

格致方法·定量研究系列　吴晓刚　主编

社会网络分析（第二版）

[美] 戴维·诺克（David Knoke）
杨松（Song Yang）　著

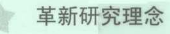
李兰　译

革新研究理念

丰富研究工具

最权威、最前沿的定量研究方法指南

30



格致出版社　上海人民出版社



格致方法·定量研究系列

1. 社会统计的数学基础
2. 理解回归假设
3. 虚拟变量回归
4. 多元回归中的交互作用
5. 回归诊断简介
6. 现代稳健回归方法
7. 固定效应回归模型
8. 用面板数据做因果分析
9. 多层次模型
10. 分位数回归模型
11. 空间回归模型
12. 删截、选择性样本及截断数据的回归模型
13. 应用logistic回归分析（第二版）
14. logit与probit：次序模型和多类别模型
15. 定序因变量的logistic回归模型
16. 对数线性模型
17. 流动表分析
18. 关联模型
19. 中介作用分析
20. 因子分析：统计方法与应用问题
21. 非递归因果模型
22. 评估不平等
23. 分析复杂调查数据（第二版）
24. 分析重复调查数据
25. 世代分析（第二版）
26. 纵贯研究（第二版）
27. 多元时间序列模型
28. 潜变量增长曲线模型
29. 缺失数据
30. 社会网络分析（第二版）
31. 广义线性模型导论
32. 基于行动者的模型
33. 基于布尔代数的比较法导论
34. 微分方程：一种建模方法
35. 模糊集合理论在社会科学中的应用
36. 图形代数
37. 项目功能差异

上架建议：社会研究方法

ISBN 978-7-5432-2171-0



定价：18.00元

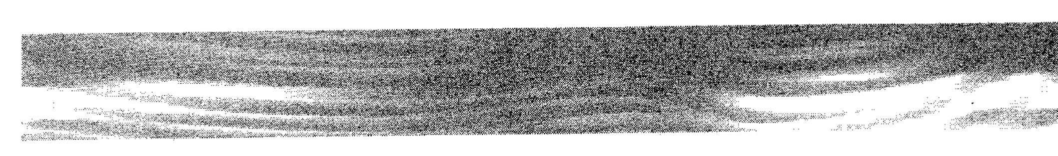
易文网：www.ewen.cc

格致网：www.hibooks.cn

格致方法·定量研究系列　吴晓刚　主编

社会网络分析（第二版）

[美] 戴维·诺克（David Knoke）
杨松（Song Yang）　著
李　兰　译



SAGE Publications ,Inc.

格致出版社　上海人民出版社

图书在版编目(CIP)数据

社会网络分析:第2版/(美)诺克(Knoke, D.),
杨松著;李兰译. —上海:格致出版社:上海人民出版社,2012
(格致方法·定量研究系列)
ISBN 978-7-5432-2171-0

Ⅰ.①社… Ⅱ.①诺… ②杨… ③李… Ⅲ.①社会关
系-研究 Ⅳ.①C912.3

中国版本图书馆CIP数据核字(2012)第216985号

责任编辑　王亚丽

格致方法·定量研究系列
社会网络分析(第二版)
[美]戴维·诺克 杨松 著
李兰 译

出　版 世纪出版集团 www.ewen.cc
格致出版社 www.hibooks.cn
上海人民出版社
(200001 上海福建中路193号24层)

编辑部热线 021-63914988
市场部热线 021-63914081

发　行 世纪出版集团发行中心
印　刷 浙江临安曙光印务有限公司
开　本 920×1168毫米 1/32
印　张 7
字　数 137,000
版　次 2012年10月第1版
印　次 2012年10月第1次印刷
ISBN 978-7-5432-2171-0/C·90
定　价 18.00元

出版说明

由香港科技大学社会科学部吴晓刚教授主编的“格致方法·定量研究系列”丛书，精选了世界著名的 SAGE 出版社定量社会科学研究丛书中的 35 种，翻译成中文，集结成八册，于 2011 年出版。这八册书分别是：《线性回归分析基础》、《高级回归分析》、《广义线性模型》、《纵贯数据分析》、《因果关系模型》、《社会科学中的数理基础及应用》、《数据分析方法五种》和《列表数据分析》。这套丛书自出版以来，受到广大读者特别是年轻一代社会科学工作者的欢迎，他们针对丛书的内容和翻译都提出了很多中肯的建议。我们对此表示衷心的感谢。

基于读者的热烈反馈，同时也为了向广大读者提供更多的方便和选择，我们将该丛书以单行本的形式再次出版发行。在此过程中，主编和译者对已出版的书做了必要的修订和校正，还新增加了两个品种。此外，曾东林、许多多、范新光、李忠路协助主编参加了校订。今后我们将继续与 SAGE 出版社合作，陆续推出新的品种。我们希望本丛书单行本的出版能为推动国内社会科学定量研究的教学和研究作出一点贡献。

总序

往事如烟，光阴如梭。转眼间，出国已然十年有余。1996年赴美留学，最初选择的主攻方向是比较历史社会学，研究的兴趣是中国的制度变迁问题。以我以前在国内所受的学术训练，基本是看不上定量研究的。一方面，我们倾向于研究大问题，不喜欢纠缠于细枝末节。国内一位老师的话给我的印象很深，大致是说：如果你看到一堵墙就要倒了，还用得着纠缠于那堵墙的倾斜角度究竟是几度吗？所以，很多研究都是大而化之，只要说得通即可。另一方面，国内（十年前）的统计教学，总的来说与社会研究中的实际问题是相脱节的。结果是，很多原先对定量研究感兴趣的学生在学完统计之后，依旧无从下手，逐渐失去了对定量研究的兴趣。

我所就读的美国加州大学洛杉矶分校社会学系，在定量研究方面有着系统的博士训练课程。不论研究兴趣是定量还是定性的，所有的研究生第一年的头两个学期必须修两门中级统计课，最后一个学期的系列课程则是简单介绍线性回归以外的其他统计方法，是选修课。希望进一步学习定量研

究方法的可以在第二年修读另外一个三学期的系列课程,其中头两门课叫“调查数据分析”,第三门叫“研究设计”。除此以外,还有如“定类数据分析”、“人口学方法与技术”、“事件史分析”、“多层线性模型”等专门课程供学生选修。该学校的统计系、心理系、教育系、经济系也有一批蜚声国际的学者,提供不同的、更加专业化的课程供学生选修。2001 年完成博士学业之后,我又受安德鲁·梅隆基金会资助,在世界定量社会科学研究的重镇密歇根大学从事两年的博士后研究,其间旁听谢宇教授为博士生讲授的统计课程,并参与该校社会研究院(Institute for Social Research)定量社会研究方法项目的一些讨论会,受益良多。

2003 年,我赴港工作,在香港科技大学社会科学部,教授研究生的两门核心定量方法课程。香港科技大学社会科学部自创建以来,非常重视社会科学研究方法论的训练。我开设的第一门课“社会科学里的统计学”(Statistics for Social Science)为所有研究型硕士生和博士生的必修课,而第二门课“社会科学中的定量分析”为博士生的必修课(事实上,大部分硕士生在修完第一门课后都会继续选修第二门课)。我在讲授这两门课的时候,根据社会科学研究生的数理基础比较薄弱的特点,尽量避免复杂的数学公式推导,而用具体的例子,结合语言和图形,帮助学生理解统计的基本概念和模型。课程的重点放在如何应用定量分析模型研究社会实际问题上,即社会研究者主要为定量统计方法的“消费者”而非“生产者”。作为“消费者”,学完这些课程后,我们一方面能够读懂、欣赏和评价别人在同行评议的刊物上发表的定量研究的文章;另一方面,也能在自己的研究中运用这些成熟的

方法论技术。

上述两门课的内容，尽管在线性回归模型的内容上有少量重复，但各有侧重。“社会科学里的统计学”（Statistics for Social Science）从介绍最基本的社会研究方法论和统计学原理开始，到多元线性回归模型结束，内容涵盖了描述性统计的基本方法、统计推论的原理、假设检验、列联表分析、方差和协方差分析、简单线性回归模型、多元线性回归模型，以及线性回归模型的假设和模型诊断。“社会科学中的定量分析”则介绍在经典线性回归模型的假设不成立的情况下的一些模型和方法，将重点放在因变量为定类数据的分析模型上，包括两分类的 logistic 回归模型、多分类 logistic 回归模型、定序 logistic 回归模型、条件 logistic 回归模型、多维列联表的对数线性和对数乘积模型、有关删节数据的模型、纵贯数据的分析模型，包括追踪研究和事件史的分析方法。这些模型在社会科学研究中有着更加广泛的应用。

修读过这些课程的香港科技大学的研究生，一直鼓励和支持我将两门课的讲稿结集出版，并帮助我将原来的英文课程讲稿译成了中文。但是，由于种种原因，这两本书拖了四年多还没有完成。世界著名的出版社 SAGE 的“定量社会科学研究”丛书闻名遐迩，每本书都写得通俗易懂。中山大学马骏教授向格致出版社何元龙社长推荐了这套书，当格致出版社向我提出从这套丛书中精选一批翻译，以飨中文读者时，我非常支持这个想法，因为这从某种程度上弥补了我的教科书未能出版的遗憾。

翻译是一件吃力不讨好的事。不但要有对中英文两种

语言的精准把握能力,还要有对实质内容有较深的理解能力,而这套丛书涵盖的又恰恰是社会科学中技术性非常强的内容,只有语言能力是远远不能胜任的。在短短的一年时间里,我们组织了来自中国内地及港台地区的二十几位研究生参与了这项工程,他们目前大部分是香港科技大学的硕士和博士研究生,受过严格的社会科学统计方法的训练,也有来自美国等地对定量研究感兴趣的博士研究生。他们是:

香港科技大学社会科学部博士研究生蒋勤、李骏、盛智明、叶华、张卓妮、郑冰岛,硕士研究生贺光烨、李兰、林毓玲、肖东亮、辛济云、於嘉、余珊珊,应用社会经济研究中心研究员李俊秀;香港大学教育学院博士研究生洪岩璧;北京大学社会学系博士研究生李丁、赵亮员;中国人民大学人口学系讲师巫锡炜;中国台湾"中央"研究院社会学所助理研究员林宗弘;南京师范大学心理学系副教授陈陈;美国北卡罗来纳大学教堂山分校社会学系博士候选人姜念涛;美国加州大学洛杉矶分校社会学系博士研究生宋曦。

关于每一位译者的学术背景,书中相关部分都有简单的介绍。尽管每本书因本身内容和译者的行文风格有所差异,校对也未免挂一漏万,术语的标准译法方面还有很大的改进空间,但所有的参与者都做了最大的努力,在繁忙的学习和研究之余,在不到一年的时间内,完成了三十五本书、超过百万字的翻译任务。李骏、叶华、张卓妮、贺光烨、宋曦、於嘉、郑冰岛和林宗弘除了承担自己的翻译任务之外,还在初稿校对方面付出了大量的劳动。香港科技大学霍英东南沙研究院的工作人员曾东林,协助我通读了全稿,在此

我也致以诚挚的谢意。有些作者，如香港科技大学黄善国教授、美国约翰·霍普金斯大学郝令昕教授，也参与了审校工作。

我们希望本丛书的出版，能为建设国内社会科学定量研究的扎实学风作出一点贡献。

吴晓刚

于香港九龙清水湾

序

社会网络分析从一开始就是一种跨学科的研究方法。与其他源于单一学科的研究方法如多维度测量法不同，网络分析的起源至少能追溯到三种学科：心理学、人类学和社会学。在心理学领域，《社会计量法》(*Seciometry*)杂志(1937)的创始人、心理医生雅各布·莫雷诺(Jacob Moreno)提出了同名法。同名法是用来测量社会关系的，以便更好地研究社会结构和心理健康的关系。人类学家 W. 劳埃德·沃纳(W. Lloyd Warner)和心理学家埃尔顿·梅奥(Elton Mayo)主持研究芝加哥西部电力公司的霍桑电厂时，罗斯里斯伯格(Roethlisberger)和迪克森(Dickson)利用这项研究于 1939 年发表了有关触排布线室的数据，即使到今天，这种分析仍然不失为一项非常出色的研究。对于这套数据，乔治·霍曼(George Homans，1950)、罗纳德·布雷格(Ronald Breiger)和他的同事(1975)利用社会学方法重新进行了分析，以使更多人有机会了解这套数据。包括哈里森·怀特(Harrison White)和马克·格兰诺维特(Mar Granovetter)在内的社会学家于 20 世纪六七十年代对推动理解社会网络、完善社会

网络分析方法，亦作出了重要的贡献。

本书第一版于1982年出版，当时书名为《网络分析》，作者为网络分析文献专家戴维·诺克(David Knoke)和詹姆斯·库克林斯基(James Kuklinski)。从那时起，社会网络分析大步向前发展，影响了社会科学的很多领域。林顿·费里曼(Linton Freeman)于1979年创办了《社会网络：一本结构分析的国际杂志》，成为该领域的旗舰杂志。许多国内或国际的社会调查都按常规把社会网络分析内容当做调查的一部分，甚至多次重复调查这些问题。例如，美国1985年的综合社会调查(GSS)收集的网络调查数据被广泛用于各种社会科学文献中。从1994年到2002年，美国做的三次全国青少年健康纵向调查都收集了社会网络方面的数据。在新千年里，欧洲的研究者首次做了欧洲生活质量的调查，其内容包括家庭、工作和生活网络等。近年来，有关社会网络分析的计算机软件如雨后春笋般涌现，而且都可以在网上免费获得。社会网络研究人员现在至少有50多种程序可以选用，例如用于社会网络分析、社会计量和序列分析的Agna，以及汤姆·斯尼德(Tom Snijder)用于分析0/1矩阵的程序ZO。

面对过去二十多年的这些发展，本书的第一版显然已经过时，亟须更新再版。戴维·诺克及合著者杨松(Song Yang)担起此任，重写了本书，现命名为《社会网络分析》。和第一版相同，第二版先是介绍了一些基本概念和数据收集问题，然后开始介绍网络分析的方法和模型。然而，即使在基本概念这部分，新版本也添加了很多重要内容，如讨论网络分析的三个主要的潜在假设。而且，这本书还反映了社会网

络分析的许多最新进展，例如，更多地强调图形及 p* 模型这样的高级方法。网络研究人员及学生将会发现，本书为他们探究社会网络数据提供了许多极好的指导。

廖福挺

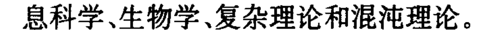
目 录

第1章

社会网络分析简介

近几十年来,社会网络越来越受到公众和学界的青睐。约翰·桂尔(John Guare)的电影作品《六度分隔》(*Six Degree of Separation*)使得社会心理学家斯坦利·米尔格雷姆(Stanley Milgrom)的小世界理论众所周知。小世界理论指的是,任何两个人都可以通过几个紧密连接的中介产生关联。该理论的另一种流行文化表现形式是"凯文·培根游戏",即影迷只需小小的努力,就可以通过演员的搭档及搭档的搭档让两个成名演员通过几个搭档联系起来。现在,在高中生和大学生中,MySpace 和 Facebook 等在线网络交流如雨后春笋般涌现;与此同时,商务部门也不断建议利用网络来寻求新的商业利润和职业发展机遇(Cross, Liedtka & Weiss, 2005)。

如图 1.1 所示,在过去 30 年里,关键词为"社会网络"的社会科学出版物成指数增长,且毫无变缓的趋势。类似地,网络研究也迅速被应用到自然现象及数学和物理领域中复杂的社会系统中(Barabási, 2002; Newman, 2003; Watts, 2003)。网络分析已然成为一种制度化的跨学科的视角,它的基本概念和测量方法已为多个领域的研究人员所熟知,如社会学、人类学、经济学、组织研究、工商管理、公共卫生、信


息科学、生物学、复杂理论和混沌理论。

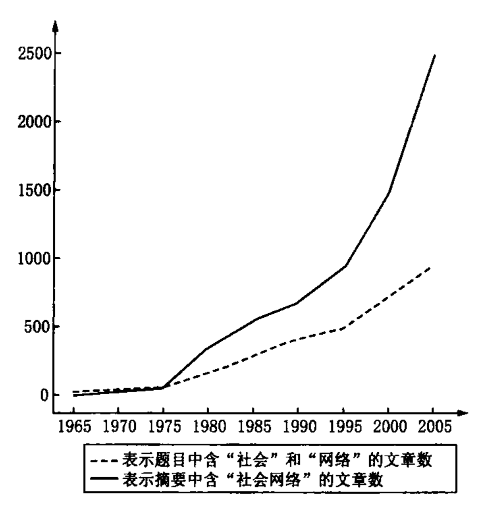


图 1.1 社会科学文献中有关"社会网络"研究文章的增长情况

同时,可用于网络分析的计算机软件也相应地快速增长,如 UCINET、Pajek、NetMiner、STRUCTURE、MultiNet 和 StOCNET 等通用程序被广泛应用于常规的网络数据分析(详细介绍请参见 Huisman & van Duijn, 2005)。同时,也出现了许多有特殊用途的程序来处理识别亚群体(Frank, 2003)、研究隐性人群(Frank & Snijders, 1994)、统计检验(Tsuji, 1997)和计算赋值图中的最优连接(Yang & Hexmoor, 2004)等特殊问题。UCINET 可能是最受欢迎、使用范围最广的软件包,版本不断更新,为许多网络分析方法的实现提供了综合的解决方案。它还提供了各种文件输入和输出、数据操作、网络工具(MDS,聚类和统计分析)和网络方法(位置、角色、亚群体、块模型和自我中心分析法)等级联模块。基于上述理由,本书将使用 UCINET 软件来分析不同的

网络实例、阐释有关的网络分析。

本书是1982年诺克和库克林斯基合著的《网络分析》(第一版)的全面修订版。除了综述有关基本方法的论题外,我们还囊括了过去1/4个世纪的研究新进展。但是,由于篇幅限制,本书不可能像沃瑟曼和浮士德(Wasserman & Faust, 1994)那样进行全面介绍,但与德热纳和福斯(Degenne & Forse, 1999)、斯科特(Scott, 2000)的入门书籍相比,本书介绍了更多的高级方法。本书为进一步学习卡林顿、斯科特和沃瑟曼(Carrington, Scott & Wasserman, 2005)介绍的更深入的高级方法打下了一定的基础。

本书主要应用于教学,针对的读者主要是社会科学领域的研究生和专业人士,其中包括社会学、工商管理、人类学、经济学、公共卫生和人力资源等。这本书可作为研究生课程的教学参考书,可用于专业学会或暑期研修班的研讨会,当然也可用于网络研究的自学。大学高年级学生和研究生如果对社会网络分析感兴趣,也可通过阅读本书在社会网络分析技巧和学术志向方面获得更高的起点。很多专业人员在做社会网络研究时会面临许多挑战,例如如何设计一个社会网络项目,网络数据收集过程中可能出现的各种细节和问题,以及分析社会网络数据时其他可以替代的技术等。社会网络学者可以把本书作为一本简单实用的复习用书或参考书。

在本书中,我们采用实际的社会网络研究问题来解释相关的概念和方法,引用的例子包括孩童游戏小组、组织、社区和跨国集团等。我们尽量避免使用复杂的专业术语,而是用简洁的语言来进行精确介绍。本书共包括四个部分。第2

章介绍有关关系数据和网络的概念，这些概念可应用于一系列的观察单元、分析层次和测量类型。这一章比较了关系的内涵和构成，区分了自我中心网络和全幅网络。结构方法强调网络分析揭示经验交流背后更深层模式的价值。第 3 章主要关注数据收集问题：边界界定、全幅网络和网络抽样、提名法和提名清单、偏差和可靠性、自我中心网络和全幅网络的关系测量，以及缺失数据。第 4 章讨论网络分析的基本方法，包括图表法和矩阵法；中心度、声望和权力；小团体、内聚和连接；结构等价；视觉演示、聚类和多维标度；块模型。第 5 章介绍了一些更高级的网络分析方法，包括位置和角色；自形等价、同构等价和正规等价；logit 模型（p* ）；隶属网络和点阵分析。读者如果希望通过分析书中的数据，以便学习如何开始进行网络分析，可以在本书作者诺克的主页 http://www.soc.umn.edu/～knoke/上找到书中所有实例的数据集和相应的计算机指令的链接，并且可以下载文件压缩包。



第2章

网络基础

本章讨论社会网络分析方法的基本概念。我们使用学者普遍使用的术语和定义,但在存有分歧的情况下,遵从社会学惯例。我们在解释这些基本概念的同时,说明这些概念的文献出处。感兴趣的读者可以通过这些文献,更好地理解网络分析方法是如何应用到实际问题的研究之中的。

为了更好地理解我们所展示的社会网络分析视角的独特性,这里将基于变量的个体分析方法与之进行一下比较。社会科学的许多领域,甚至社会科学的绝大部分领域,都假定行为人作出决定和采取行动时,不考虑其他行为人的行为。无论是基于效用最大化的理性计算来分析,还是基于因果关系的动因推理来分析,这些解释仅仅主要考虑了个人特性,而忽略了社会行为人所嵌入的更广泛的互动背景。与此不同的是,网络分析法明确假定,行为人参与到社会系统之中,而这个系统将他们与其他行为人联系在一起;行为人之间的关系对各自的行为存在着重要影响。网络分析理论和方法的关键,是识别、测量和检验关于行为人之间关系的结构形式和实质内涵的假设。这种对结构—关系研究的独特重心,将社会网络分析法与迄今依然主宰大部分社会科学的以变量为中心的个体分析的传统区分开来。

第1节 基本假定

网络分析的视角强调把结构关系作为关键的导向原则，其中社会结构的组成是“有形实体之间关系模式的规律性，而不是抽象的规范和价值之间的协调，也不是对这些有形实体的特征进行的分类”(White, Boorman & Breiger, 1976: 733—734)。实体可以是单个的自然人、小的群体、组织甚至是国家。通常连接这一系列实体的关系模式包括宏观的社会背景或者整体的结构，这些背景和结构影响它们的认知、信念、决定和行动。网络分析的首要目标是精确测量和展现这些结构关系，并解释他们发生的原因及产生的结果。

社会网络分析的重要性依赖有关关系模式及其效用的三个基本假设。首先，对于理解所观察的行为，结构性关系要比年龄、性别、价值和意识形态等个体特征更为重要。例如，在解释15世纪的佛罗伦萨哪些精英家族支持美第奇(Medici)或寡头政治集团的统治时，更多地依赖这些统治精英之间的经济、赞助和婚姻关系，而较少依赖他们的阶级和社会地位等个人特征(Padgett & Ansell, 1993)。在实体参与的不同社会情境中，他们的许多特征保持不变(如不论在家、在工作场所还是在教堂，一个人的性别、种族和教育程度都是不变的)。但是，具体的结构关系仅存于特定的时空；当

情境发生变化时，这种特定的结构关系将消失或暂时中止(例如，学校之外就不存在老师和学生之间的关系；当一个人丧偶或离婚时，婚姻关系将不存在)。一个在办公室做勤杂工的女性，在工作场所可能不需要什么主动性，但她却可能是邻居关系协会中强有力的领导者，或者是个敢作敢为的PTA[①]成员。这种行为差异很难用性别、年龄和社会地位等不变的特征来解释；但是，当意识到人与人之间的结构关系可以随着社会情境的改变而变化时，这种差异就很好理解了。这种结构关系的分析方法备受网络分析者的欢迎；而与之形成鲜明对比的是诸多的互争高低的实体论方法，这些方法将静态的“事物—概念”作为自己的主要分析单元，包括本质、自我行为、基于规范的一致、理性选择、以变量为中心和社会认同途径，等等(Emirbayer，1997)。网络分析认为，除了个体特征，社会实体间的已具有模式的关系也影响社会实体。根据这样的推测，网络分析宣称，可以为分析社会行为的原因提供更加全面的理论和实证解释。

第二，社会网络通过实体之间的关系所建立的各种结构机制影响实体的观念、信仰和行动。直接关系和更密集的互动关系使得实体获得更好的信息、更多的认知、更强的敏感性，以影响他人或受他人的影响。间接关系通过媒介(例如常见的经纪人，他们为客户牵线)也会激发新想法，或获得一些潜在的有用资源，而这些资源可能需要通过与他人交易才能获得。例如，相对于较弱、较疏远的社会关系，求职人员一般很难通过更密切的社会关系获得有价值的信息，因为关系

① 即家长—教师关系协会。——译者注

更亲近的人已经共享或传播了相同的信息(Granovetter, 1973)。关系结构提供了各种复杂路径以协助或阻止人群中知识、流言蜚语的传播。社会结构配置上的不同,解释了芬兰有机耕作的普及(Nyblom, Borgatti, Roslakka & Salo, 2003)和美国城市中性传染疾病的传播(Jolly, Muth, Wylie & Potteral, 2001)。个人的身心健康相当大程度上依赖社会支持网络——比如,当你沮丧的时候,是否有人给你送上一份心灵鸡汤(Cohen et al., 1998)。结构关系对维持团体的凝聚力和团结至关重要,但也可能强化对外部群体的偏见和冲突。竞争与合作关系有效实现了跨国公司内部有计划的组织变迁(Tenkasi & Chesmore, 2003)、社会运动中集体行动的动员(Diani & McAdam, 2003),以及从事贩毒、非法移民、走私和恐怖活动等"黑色网络"的运转(Raab & Milward, 2003)。通过把信息和资源传递到特定的结构位置,网络有助于激发兴趣、分享认同、促进共同的价值和规范。网络分析者试图揭示社会关系如何影响社会实体的理论机制,并发掘出特定机制在具体情境中运行可能需要的条件。

网络分析的第三个潜在假设是应将结构关系视为动态过程。这个假设认识到网络不是静态结构,而是通过个人、群体或组织等实体内部的互动不断改变的。当这些实体利用有关网络的知识获取优势时,他们或有意或无意地改变了他们所处的关系结构。例如,在柯尼斯和诺克(Kenis & Knoke, 2002)有关组织领域网络变迁的理论中,就认为先前的沟通结构会影响后续的战略联盟抉择。相应地,这些组织间的关系改变了信息的流动,从而为将来的联盟提供了更多的机会或约束。这些动态过程验证了社会行动理论中更为普遍的

"从微观到宏观的问题"(Coleman, 1986)。关键问题是大规模的系统性变化如何从个人的综合偏好和目的性行动中显露出来。由于网络分析同时涵盖实体和结构,因此它为微观层面选择的变化和宏观层面结构的变迁之间的连接提供了有效的分析概念和方法。不幸的是,跨层动态过程的实证网络分析仍只是一个理想的目标,并没有成为一个普遍的做法(Emirbayer, 1997:305)。但是,随着纵向网络数据的逐渐增多,对网络变迁进行抽象和调查的新方法也将出现(Faust & Skvoretz, 2002; Snijders, Steglich, Schweinberger & Huisman, 2007)。

第2节 | 行为人和关系

任何社会网络中两个不可或缺的元素是行为人和关系。它们的结合共同构成了我们下节要介绍的社会网络。行为人可以是自然人等个体,或非正式群体、正式组织等集体。个体行为人的例子包括操场上的小孩,参加舞会的高中生,公司里工作团队的职员,疗养院的员工和住客,秘密基层组织中行动的恐怖分子,等等。集体行为人可能是在一个行业中竞争的公司,筹集慈善基金的志愿者组织,在国会里拥有席位的政党,也可能是签署军事同盟协议的国家。有时网络行为人也包括一些混合类型,如一个组织领域包含有供货商、生产商、顾客和公共卫生的政府监管人员。

关系一般被定义为两个行为人或对点之间的特定接触、连接或联结。关系可以是定向的,即第一个行为人发起,第二个行为人接受(如劝告);关系也可以是非定向的,即存在互动(如交谈)。关系不是一个行为人的特征,而是在只有两个行为人保持联系时一种共同的对偶特征。在个体行为人以及集体行为人之间,会发生许多各种不同的关系,这些关系可展示网络结构,并解释它们的效应。比如,在个人之间交往的层面上,小孩之间交朋友、一起玩耍、打架、相互信任,

同事之间一起工作、讨论、提意见、信任、背后说坏话和背叛。而在集体行为人之间,公司进行商品和服务的交易、相互交流、竞争、起诉、游说和合作。一个网络研究者应该测量哪一种关系,这取决于研究项目的特定目标。例如,社区网络研究可能调查各种不同的邻里活动,而银行网络研究则主要调查财务交易。当然,一些分析详细研究多类关系,如公司董事会董事之间的政治、社会、经济联系。下节我们将介绍关系内涵的一般分类。

很多社会研究无论是通过调查还是实验来收集数据,依旧依赖测量和分析行为人的特征,并把行为人的特征作为分析单元。虽然特征和关系是研究社会行为的不同概念方法,但是它们不能视为互斥的选择。相反,许多行为人的特征可以重新概念化为两个行为人之间的关系。因此,虽然一国每年的进出口额是这个国家经济的特征,但所有的任何两个国家间交易的商品和服务量却可以表示全球经济的交易网络结构。一个棒球队的总输赢纪录显示这个球队的表现,但每两个队的输赢比例可能揭示一个联盟权利结构的细微差别。朋友的数量显示了一个小孩的受欢迎程度(人气),但分析每一对朋友关系的选择就能揭示出重要的小团体和群体。关系反映出复杂社会系统中仅靠成员特征的求和或均值所不能描述的新维度。结构关系能影响个人行为和系统表现,且这些行为和表现不能还原为行为人的特征。例如,与其考虑行为人的教育、年龄、社会阶层、性别和种族,不如考虑行为人之间的交流和建议关系结构,这样能更好地解释技术创新的传播比例和模式(Cowan & Jonard, 2004; Valente, 1995)。类似地,和推测群体的易受感染性和抵抗力特征相比,熟知群

体间的亲密关系结构更有利于防止像禽流感和艾滋病等传染性疾病的流行爆发(Jolly et al.，2001)。在此，一个非常重要的推论是：仅仅分析行为人的特征，会忽略社会行为的网络视角所提供的许多重要解释。

第3节 | 网络

社会网络是由一群行为人组成的结构，在这个结构中，行为人通过一系列关系相连。社会网络的绝大多数定义都包括这两个基本组成部分。例如，“社会结构可以表述为网络——包括一系列节点(或社会系统的成员)和一系列描述它们之间关联性的关系”(Wellman & Berkowitz, 1988:4)。即使行为人都相同，不同类型的关系也可以形成不同网络。因此，一群雇员之间的朋友网络很可能与他们之间的建议寻求网络存在很大差异。网络行为人之间存在联系并不意味着所有行为人必须和其他行为人之间存在直接关系。实际上，有时直接的对偶关系并不多见。相反，网络分析会同时考虑目前存在的或不存在的关系，也可能会把关系的强度变化考虑进去。有形实体之间的实际关系构造，便确定了一个特定的网络结构——该网络的模式或形式。网络结构的形式变化多端，既有所有行为人之间没有关联的孤立结构，也有所有行为人之间直接相关的饱和结构。一般来说，实际的网络结构处于独立网络结构和饱和网络结构之间，其中一些行为人比其他行为人拥有更广的网络。网络分析中的一个核心理论问题，便是解释不同结构的发生情况，以及在行为人层面来说明其与其他行为人之间关系的变化。在网络研

究中,一个同样重要的实际任务是利用关系数据,发掘并精确展现这些结构。

约翰·A.巴恩斯(John A. Barnes, 1954)被认为是第一个使用“社会网络”这一术语的学者。他是一个人类学家,曾经研究挪威一个岛上居民之间的关系。巴恩斯将社会互动视为“点的集合,其中一些点由线连接”,从而形成关系的“全部网络”(Barnes, 1954:43)。人际关系的非正式形式构成社会网络整体的一个部分。巴恩斯利用了雅各布·莫雷诺(1934)的成果,后者在展示小孩对同班同学的爱憎关系研究中,率先使用了线和标示点的手绘社会关系图。本书将在第3章讨论如何表示有关社会网络关系的形象图示和数学矩阵。半个多世纪以来,社会网络这一理念和方法已经从人类学和社会学延伸到诸多学科,并被运用到诸多主流的理论和问题研究中。关于社会网络原理缘起及普及的历史性综述,可参见费里曼(2004),斯科特(2000)及诺克斯、萨维奇和哈维(Knox, Savage & Harvey, 2006)的著作。

假如社会网络分析仅仅只是描述一群行为人是如何联系的一个概念性框架,那么它就不会激起社会研究者如此大的兴趣和努力。作为一套综合的理论概念和分析方法,社会网络分析所能提供的不仅仅是精确重述。由于社会网络结构影响到个体层次分析和系统层次分析,因此它被认为可以用来解释结构关系中的变化及其结果。J.克莱德·米切尔(J. Clyde Mitchell, 1969:2)关于社会网络的定义,就特别强调了它们对结果的影响。他的定义如下:社会网络是“在特定人群中的一整套具体的联系;这种联系有一个额外特征,即这些联系作为一个整体可以用来解释其中人们的社会行

为”。这一论述在本书的第一版中也得到了回应:“行为人之间的关系结构和个体行为人在网络中的定位,对于每个个体单元和整个系统来说,在行为、意识和态度方面都有重要的影响。”(Knoke & Kuklinksi, 1982:13)同样,巴里·威尔曼(Barry Wellman, 1999:94)也写道:“在描述社会结构的基本模式,在解释这些模式对人们行为和态度的影响上,社会网络分析都能起到作用。”

第4节 | 研究设计的基本要素

对于研究者来说，构成测量和分析框架的社会网络研究设计的三个基本要素是：抽样单位、关系的形式和内容、数据分析层次。每一个网络研究项目都必须在开始田野工作之前对每个基本要素作出明确界定。这些设计因素的不同组合将生成各种不同的社会网络调查，这将都呈现在多个学科的研究文献中。

抽样单位

一项社会网络研究设计的第一步，即选择一个最相关的社会背景，并确定在此背景下的哪些组织涵盖了这些社会网络行为人。根据规模和复杂性大致的升序排列，构成抽样单位的六个基本单元，包括个体、群体(正式和非正式)、复杂的正式组织、阶级与阶层、社区，以及民族—国家。一些两阶段的研究设计包括一个高层次系统，在此系统中行为人构成其中的低层次实体。这种嵌套社会背景的普通例子包括公司与职员、学校与学生、市政机构与公务员、大学及其院系与教授等。

最早且至今仍然流行的社会网络研究项目一般选择小

规模的社会背景——如课堂、办公室、工厂、帮会、社交俱乐部、学校、村庄、人为组织的试验群体等,在此将其中的个体成员作为社会关系的行为人,而他们的关系正好构成了受调查的网络。最新的例子有:加州少年课堂中与欺负和被欺负相关的朋友关系模式(Mouttapa, Valente & Gallaher, 2004),两家冶金工厂全方位质量管理方案的实现(Harrison, Laplante & St-Cyr, 2001),以及马达加斯加一个村落中家庭生育控制方式的社区分布情况(Stoebenau & Valente, 2003)。在精确描述成员资格、全面观察群体、易于通过上层权威打通关系等方面,小规模背景都有明显优势。较大规模的环境往往都有互相渗透的或者模糊不清的边界,但是社会网络分析法并不因其内在特性而不适宜将其概念与方法运用于这类环境。最近的研究范例有:某一中西部城市青少年的恋爱与性行为关系网络(Bearman, Moody & Stovel, 2004),学院派经济学家的合著关系(Fafchamps, van der Leij & Goyal, 2006)。

关系的形式和内容

除了选择一个恰当的抽样单元之外,社会网络研究者还必须决定收集哪种具体社会关系的数据。行为人之间的关系既有形式又有内容,这种两分法概念是由格奥尔格·齐美尔(Georg Simmel, 1908)在分析社会交往时提出的。形式和内容这二者在实践中是不可分的,只是为了分析的方便而将二者区分开来。内容是指社会交往中人们的兴趣、目的、压力或者动力,而形式则是由此而将具体内容转化为社会现实的互动模式。齐美尔认为,社会学的任务就是发掘确认社会

交往有限的几种形式:交际、超越、臣服、竞争、冲突、合作、统一等。这些交往形式是在广泛的具体环境、社会制度和历史背景下发生的。同一种具体形式在内容上可以千差万别。例如,上级和下级这一基本关系形式,可以在政府、军队、商业、宗教、体育运动和文化机构等不同组织中展现。相反,不同的内容——比如追逐经济利益和政治权力——也可以通过竞争与合作的形式展现。

内容与形式这一二分法概念也运用到社会网络分析中。关系形式是行为人关系的一个特征,它独立存在于所有的具体内容之中。有两种基本的关系形式:(1)双方行为人之间互动关系的强度、频率或者力度;(2)每对成员之间关系的指向,即是否为0,或者非对称的,或者双向的选择。关系内容指作为关系发生原因的那些事物(Burt, 1983:36)。这种实质性内容是由研究者建立的一种可分析构造,意在从行为人主体视角获取某一关系的意义。当人们在某一社会背景中被问及"你的亲密朋友、一般朋友和认识的人"时,这一问题的内容就是指"友情"。这一询问的结果需要依每一行为人是如何定义这三个词所代表范畴的意思而定,另外还得依行为人是如何根据不同人际互动的以往经验来区别对方而定。很明显,人们在对友情这一标签及表示亲密或疏远的行为的理解上,可能会有很大差别。

关系内容或者说连接模式的选择,大部分是由项目的理论焦点和研究目标来决定的。因此,一项关于保健关系网络的研究就会调查人们有关健康事项信息和建议的来源,而一项关于政治关系网络的研究也许会让他们列出平时经常和他们一起讨论或参与政治事务的人。在一些实证问题的研

究中,也许不仅仅某一个独特的可分析关系内容需要调查,而且其他两三种甚至更多的连接模式(即复杂关系网络)同时也可能需要进行测量和分析。一项关于德国、日本与美国政府劳工政策游说机构的比较研究,要求受访人报告他们的独立政策信息、政治支持、与他人的信息互换关系等(Knoke, Pappi, Broadbent & Tsujinaka, 1996)。我们将在第 4 章中利用其中的一些数据来阐述有关网络分析的方法。

在齐美尔看来,对于将关系内容的细微差别进行概念化和操作化,虽然研究者们几乎有着无限的想象力,但是在这里我们还是提供一个小小的基本内容范式,具体内容如下。

(1) 交换关系:行为人互换有关实物或象征性物质的控制,比如赠送礼物或者买卖商品。

(2) 沟通关系:通过行为人之间的联络渠道,双方交换有关信息。

(3) 交叉关系:行为人在两个或多个社会组织中有成员资格,例如公司董事会成员同时拥有其他身份。

(4) 利用关系:行为人之间互相联系,是为了获得有价值的物品、服务或者信息,比如一份工作、一个政策建议或者一个加入某项社会行动的机会。

(5) 情感关系:也许最常接受调查的社会网络关系就是行为人对他人的感动、崇拜、尊敬、厌恶或者敌视等情绪的表达。

(6) 权力关系:这类联系通常发生在正式的科层组织中,表明行为人对于发布和遵守命令的权利和义务。

(7) 亲属关系:这类血缘和婚姻的联系反映家庭角色内的关系。

社会网络分析对于各种各样的关系内容是如何连接的仍然认识不足,因为对于关系内容领域结构的考察仍然研究得不够。罗纳德·伯特(Ronald Burt, 1983)在关于调查对象是如何认识关系内容的研究中,发现在抽样调查北加州人的33个问题中,存在实质性的混淆、重复和可替代等情况。他认为,仅仅只需五个关键问题就足以涵盖有关友情、相识、工作、亲属和亲密关系这几种关系内容的主要结构。然而,对于在各种不同网络环境中普遍使用的关系术语和标签,人们所赋予其意义的相似性和差异性,我们还需要更多的研究。一张关于关系内容领域间的结构性连接的认知图,可以帮助研究者准确有效地选择最适合理论和实践关注点的具体关系内容。

数据分析层次

在选择了抽样单元、关系的形式和内容之后,研究者对于他们在社会网络研究项目中所搜集到的数据,在哪一个层次上进行结构分析,有多个可替代性选择。本书第3章和第4章会详细论述有关合适的测量和分析方法。这里我们将首先概述四种概念独立的分析层次,调查者可能会从中选择决定集中分析。

最简单层次是自我中心网络。这一社会网络由一个行为主体和其他行为客体组成,其中行为主体与其他行为客体发生直接联系,而其他行为客体之间也存在直接联系。这一层面也被称为行为主体的“第一关系圈”;相对于第一圈,所有行为客体的客体构成了第二关系圈和更高层次的关系圈,

以此类推。假如一个网络中有 N 个行为人，那么自我中心网络则会有 N 个分析单元。相应地，每一主体行为人可以通过其与其他行为客体联系的数量、强度和其他特征加以描述，例如报酬与付出关系的比例，或者行为客体之间联系的密度。最近一项关于毒品成瘾者个人健康风险的自我中心网络研究发现，波多黎各的瘾君子在与陌生人或认识的人发生性行为中，要比纽约的瘾君子更容易感染艾滋病病毒(Kang, Deren, Andia, Colon & Robles, 2005)。在某些方面，自我中心的分析层次类似于典型的基于属性调查研究，即在受调查者通常的个体特征——例如性别、年龄和教育背景等属性之外，再补充一些可从其直接网络关系中获得的测量值。自我中心网络研究设计比较适合调查那些不太愿意与他人接触的对象。1985 年美国成年人口社会综合调查(Marsden, 1987)率先识别和收集有关受调查对象的客体的信息，我们将在第 3 章详细论述。

第二个分析层次是由成对的行为人构成的对偶网络。假如双方组合顺序无关，就像在婚姻状况中两个人要么是未婚、同居、已婚，要么是分居、离婚等，那么由 N 个行为人组成的样本则有 $(N^2-N)/2$ 种分析对偶组合。但是，如果组合关系的指向有意义，比如下达命令或者征求意见，那么这一样本包含 (N^2-N) 种有序对偶组合。对偶网络最基本的问题是：两行为人之间是否存在具体联系？如果是，那么这种联系的深度、强度和持续时间又是多少？与此密切相关的一个问题是：没有直接联系的双方是否仍然经由中间渠道(比如经纪人、中间人)间接联系？一般的网络分析都试图将对偶关系的各种变化解释成二元特性的一个函数，比如羽禽群

居的同质假设,或者是异性相吸的互补假设。与一般看法相反的是,一项关于1951年至1985年间国际关系冲突的研究发现,具有紧密军事与经济联系的两国和缺乏如此紧密关系的两国相比较,在陷入争端的可能性上前者也只是比后者稍微少那么一点点(Benson, 2004)。

第三个网络分析层次当然就是三元关系了。N 个行为人有 $\binom{N}{3}$ 种组合,这是从 N 个行为人中每次取3个得到的组合数。在3人一组的行为人中,显性和隐性可选择关系的所有可能组合将产生一系列16种独特的三元组类型。在实证网络分析中,一个基本的描述性问题就是在16种类型中观察到的三元组的分布情况,这种综合分布量表也被称为三元关系统计表。三元关系结构的实证研究一般关注情感性连接点(比如喜好、友谊、敌视),并且特别关注三元关系的平衡和传递(例如,假如A选择B, B选择了C,那么A是否也会倾向于选择C呢)。因为本书没有足够篇幅来评述三元网络分析方法,有兴趣的读者可以关注詹姆斯·戴维斯、保罗·霍兰和塞缪尔·雷恩哈特(James Davis, Paul Holland & Samuel Leinhardt, 1979)的研究项目,以及沃瑟曼和浮士德(1994:556—602)的全面综述,以获得更加详细的资料。

除了上述三个微观层次的分析之外,全幅网络是最重要的宏观分析层面。研究者利用所有 N 个行为人的每一种关系的信息,对整个网络的结构关系进行描述和解释。其中,研究者特别关注在此网络系统中行为人的特定位置或者社会角色的展现,以及在这些位置上和位置间行为人联结的模式。虽然一个全幅网络有 N 个行为人和 (N^2-N) 种对偶组

合(自我关系一般被忽略),但是所有这些元素加起来只形成单个系统。在全幅网络层面上分析结构变化的原因和结果的各种假设,可能需要利用多个不同系统,而且也是相当耗费研究成本的。然而,全幅网络分析却广受欢迎,其中一些例子包括对爱尔兰都柏林地区青少年朋友关系和伴侣关系的研究(Kirke, 1996)、世界各城市间优势等级分层的研究(Smith & Timberlake, 2002),以及全球信息产业部门战略联盟的研究(Knoke,待出版)等。

网络分析的这四个层面表明,在某一个层次上涌现的现象,并不能够由其他层次上关系的现有认识简单地通过逻辑推导出来。例如,选择关系的传递在友谊构成的理论上是一个非常重要的变量(我朋友的朋友,就是我的朋友),这在三元关系层面中可以观察到,但是在自我中心关系和对偶关系层面上就不可能观测到。举另外一个例子:考察两个科学研究团体,他们在科研讨论网络中具有大概相似的自我中心结构、对偶结构和三元结构。但是,假如第一个科研团体的全幅网络分为几个互不相关的小组,那么其中的许多科学家就不可能与其他人接触,甚至间接接触也不可能。假如第二个科研团体的整个网络中各小组间有联系纽带和通道,那我们就可以想象到,其中的信息沟通更加流畅,科学创新的频率更高。网络分析作为一种实证研究的理论指导框架,它涵盖各种突发的结构关系,从而能够解决多个分析层次上的问题。网络分析这一变化多端的功能使得它迅速流行起来。

第3章

数据收集

对于任何网络实证研究，调查者在开始收集数据之前，必须首先处理三个重要的事项：范围界定、网络抽样、关系测量。由于在网络研究项目里，设定行为人的边界是一个不言而喻的起点性工作，因此我们就从这点开始讲起。

第1节 | 范围界定

关于网络分析的范围界定问题，可以这样简单地设问：在收集社会关系的数据时，也许这些社会关系实际上并不存在一个清晰的界线，那么一个研究者在哪里设定边界呢(Barnes，1979:414)？罗曼、马斯登和普伦斯基(Laumann，Marsden & Prensky，1983:19)一开始通过对比唯实论与唯名论的定义策略来寻求他们的答案。在设定具体网络行为人的资格以及确认待分析社会关系的类型上，他们制定具体的涵盖规则。但在随后的网络范围界定范式上，他们用三个基本方法取代了上述二分法策略：位置、关系和事件(Laumann，Marsden & Prensky，1989)。本节我们将分析这两种处理过程。

唯实论和唯名论的界定策略

在唯实论的界定策略中，网络分析者采纳系统行为人自身的主观意向，将范围界定在某一实体组织(例如一个家庭，一个公司或者一个社会团体)内的所有或者几乎所有的行为人都可以自觉体验到的限度内。行为人及他们的关系是包含还是排除，得依赖于其他行为人对于相关性的判断。例

如,诺克和罗曼(1982:256)即采用这一规则来选择美国能源与健康国家政策领域中的核心组织:"一系列重要的行为人旨在参与起草、建议或选择有关行动方针(即政策选择),以期解决当前明确的重要问题。通过确定他们之间的相互关联性或者共同目标的实质性定义标准,便得到了一个作为子系统的政策领域。"有些组织在国家能源和健康政策的制定上,被圈内人士认为并不具备什么影响力,因此研究者就将它们排除在外。

在唯名论的界定策略中,为某一特定研究项目的分析或理论目标服务的一个概念框架,在这里被当做一个先在的理念,用于界定某一网络边界。在许多情形下,法定的或者正式的成员资格要求已经划出了一条明确的界线,比如某一课堂上的小学生、某州律协的律师、在联合国具有普遍协商身份的组织。另一个例子是,马克思主义研究者在定义工人阶级时,将其界定为具有共同生产关系的所有雇员。当然,在何种程度上,主观评价和概念强加这两种界定策略恰好会得到相同的边界,这完全是个实践的问题。

定位策略

这种策略利用行为人的特征、他们在某一正式组织内的成员资格,或者对某一个清晰界定的位置的占有,来对某一网络进行边界划定。尤西姆(Useem, 1979)对美国商业精英的研究,便是这一定位策略的范例。他将商业精英界定为在1969年的797家美国最大公司的董事会里的成员,并从中选出了8623个样本。诺克(2001)在分析全球信息产业部门的

公司战略联合时，将公司在1989年至1998年间《财富》500强、《财富》1000强、全球500强的榜单排名作为遴选标准。费尔德曼·萨维斯伯格、东科和杨（Feldman-Savelsberg, Ndonko & Yang, 2005）调查Bamiléké城六个喀麦隆妇女同乡会成员的集体记忆选择时，要求这些妇女都来自同一个村庄或者部落。研究者利用定位策略可能会发现，某一组织的成员名单或者花名册已经过时、不完整，或不准确。他们需要自己做普查来编制一个完整的成员名单。

定位策略一般会得到一个行为人的集合，他们在某一正式组织中拥有相似的位置，即使有的人彼此之间并没有直接关系。对于使用定位方法揭示的网络结构的代表性，研究者应该保持足够的警惕。因此，商业精英间的联系与相对低阶位的雇员间的联系是大大不同的。跨国公司间形成的战略联盟与小企业之间的合伙也是大不相同的。伴随定位策略而生的另一个网络问题是，行为人之间往往并无联系，或者构成一些小的紧密团体但彼此间缺乏联系。例如，有些喀麦隆妇女，虽属同一同乡会，但是通常彼此间却完全陌生（Feldman-Savelsberg, Ndonko &Yang, 2005）。

在采用定位方法进行网络分析时，研究者应当对涵盖或排除某一特定定位标准提出明确合理的理由（Laumann et al.，1983；Scott，1991:58）。研究者可以采用唯名论标准，为他们的排除规则设定一个主观性门槛，即使某一定位处于持续变化之中。例如，为了研究大公司，研究者也许需要从《财富》榜单中抽取50家、100家甚或500家公司。如何设定选择范围，更多地取决于时间和研究经费限制，而不是取决于入围公司与未入围公司的某些实质性差别。戈拉斯克维

奇(Galaskiewicz, 1979)在对一个叫 Towertown 小城的一些组织进行筛选时,就生动地说明了上述过程。首先,他采用属地原则,即将研究限定在一个地理范围内的人口。其次,他根据他的研究时间和经费限度,利用行业标准,排除了商业机构、交通业、公用事业、地产业、街道居民组织、社区机构以及小学等。

关系策略

这一方法依赖于知情者或网络行为人本人提名添加其他行为人。关系策略包含几个常用的做法,比如声望法、滚雪球抽样法、固定名单选择、扩展性选择以及 K 核心方法。我们将在本小节简要地描述一下这些关系策略方法所需的条件及其局限性。

在声望法中,研究者通常咨询有关资深知情人士或者专家,请他们提名有关行为人。例如,调查者调查两个精神健康治疗机构网络,首先统计当地社区精神健康治疗机构总名单(Morrissey, Tausig & Lindsey, 1985)。接着,他们会询问调查对象,请他们从总名单中挑出五个他们认为最重要的,或者他们最经常联系的机构。通过这两个步骤,便挑选出了两例核心的精神健康治疗机构。接下来,请关键知情者即人道服务计划委员会和矫正部门的主任们,从总名单中再挑选额外的重要机构,对核心名单进行补充。

声望法非常依赖关键知情人提供准确全面的信息,因此研究者在确定重要知情人士上的能力便非常值得关注(Scott, 1991:59)。莫里西(Morrissey et al., 1985:35)就曾

经忠告:"在评估声望法划分标准的准确性上,并没有一个统一的准则。很显然,不同的划分标准会产生不同的行为人涵盖范围,也会导致随后的分析受到影响。"对采用声望法获得的网络关系名单的可靠性和完整性进行评估,是一件非常困难的事情。经常只有等到一个项目完成之后,才可能对此进行评价。因此,网络研究者挑选关键知情人,应当有充分的理论和实践依据,且这些理由是独立于受调查的具体社会关系的(Scott, 1991:59)。但有时抽样偏差问题——即由于非随机性的案例缺失而导致样本不能代表目标总体(Allison, 2001:78—81)——会超越关键知情人士的遴选问题。关键知情人士不管如何博学多才,他们所提供的数据,总是与通过滚雪球抽样法所得到的数据有系统性差别。

在滚雪球抽样法中,研究者从小范围的网络行为人开始,由他们来提名与其有某种特定关系的其他参与者。同样,这些其他参与者也被要求提名其他行为人,直到没有或者几乎没有其他人名再被提及(Frank, 2005; Wasserman & Fraust, 1994:34)。在此方法的早期严格版本中,每一轮滚雪球抽样都将产生等量的行为人(例如,列出三个你最好的朋友),并且在每个阶段每一个行为人都被问及同样的问题(Goodman, 1961)。最近,这两个特殊条件有所松动,允许使用不同的问卷,提出数量不等的行为人(Wasserman & Faust, 1994:34)。因为滚雪球抽样法利用网络行为人的社会关系来构建样本,所以每一轮都会提出新的行为人,而这些新的行为人与现有的行为人存在特定关系。因此,滚雪球抽样通常会得到一个关系非常紧密的社会网络(Laumann et al., 1983),因而被称为连锁推荐法(Heckathorn, 1997)。应

答驱动抽样是该方法的一个复杂延伸,它将滚雪球抽样法与某一数学模型结合,从而增加抽样的代表性,以弥补非随机性选择所带来的偏差(Salganik & Hckathorn, 2004)。

在社会网络分析中,滚雪球抽样法对寻找难以接触人群具有独特效果,比如毒品交易者、吸毒者、非法移民、HIV 阳性者的性伴侣、地下抵抗组织等。由于缺乏对此类隐蔽人群的抽样框,常规抽样调查方法经常得不到足够的样本量。相反,滚雪球抽样法可以从访谈一个初始样本开始,比如瘾君子经常光顾的治疗机构或者街区。受访者被要求提供与之有特别关系的其他行为人的地址信息。接着,调查者联系这些被提名者,并且请他们提供与其相关的其他参与者。因为网络数据收集工作常常需要知道行为主体介绍的行为客体的身份,所以获得其知情同意、保护匿名来源、确保隐私等要求就引发了一些伦理问题(Borgatti & Molina, 2005; Klovdahl, 2005)。如果研究者直接利用滚雪球抽样法来招募行为客体,人权事务委员会也许就不会批准这种抽样设计方案。作为替代办法,应答驱动抽样设计会要求种子行为主体告知他们的行为客体,如果自愿参与项目可以得到一点小小的补偿。

在实施网络抽样时,一个程序性问题——是否给知情者提供一个人员名册供其选择还是让受访人自己提名——将选择方法分成了两种:固定名单选择法和扩展选择法(Doreian & Woodard, 1992)。在固定名单选择法中,受访人被限定于在研究者和知情人士事先所确定的行为人名单中,报告其社会联系。在扩展选择法中,受访人可以不用参照某一名单而是按照自己的意愿确认其他行为人。扩展选择法的运

用非常类似于滚雪球抽样法的程序。在多利安和伍达德(Doreian & Woodard, 1992)的一项关于小孩和未成年人服务体系项目机构的研究中,任何得以涵盖在此网络关系中的服务机构必须满足下面两个条件中的一个:现任机构主管至少引用三次,或服务员工至少提名五次。低于上述提名门槛次数的机构将被淘汰出局。多利安和伍达德的报告中指出,固定名单选择和扩展性选择得到的网络结果,在几个方面都存在很大差异,其中包括行为人数量、关系节点数量、关系点的密度和质量等(关系点的密度是通过将观测到的双向关系点数除以所有可能的关系点数得出来的,关系点的质量是知情人士对关系的满意度、可延展性及价值的估量)。特别需要指出的是,固定名单法只从总行为人中选出50%的行为人,40%的双向关系节点是通过扩展性选择方式发现的。固定名单选择方法更容易产生非随机性抽样偏差,也就是说,它总是产生一套核心行为人名单,而系统性地排除了边缘关系行为人。这个方法会有个不好的结果,即得到一个没有结构性背景的关系网络。除非有条件可以确保两种方法都得到一个相差无几的网络,否则固定名单选择法并不能作为扩展选择法的替代方案,即便它有管理成本低的优势。

最后一种关系策略是K核心法。赛德曼(Seidman, 1983)提出了K核心的概念,用于发现在一个大网络中的行为人子集。这些行为人子集一般包括许多的小团体,小团体之间关系微弱,但是在团体内却团结紧密(Yang & Hexmoor, 2004)。假如在一个小团体中,每一个行为人与至少K个其他行为人存在联系,则这个小集体就是一个K核心。通过改变K值,研究者可以为某一网络的边界设定或宽或窄的

限制标准。多利安和伍达德(1994)运用扩展选择法,描述了K核心概念是如何用于定义和确定网络边界的(我们将在本书第4章更详细地讨论K核心法)。通过改变K的阈值,研究者能够使一个非常大而松散的网络边界重新得到划定。当K值高时,边界标准更严格;当K值低时,边界标准更宽松。为了实用目的,多利安和伍达德(1994)推荐使用一个低K值,以便建立一个更全面的网络边界。相对来说,一个涵盖范围更广的网络更少受到抽样偏差方面的质疑,同时研究者随后总是可以运用高K值来构建一个更加严格紧密的关系网络。然而,一个涵盖范围大的网络却会带来更重的数据收集任务:虽然低K值会使得网络范围更宽,但是研究者在每一轮都必须访谈数量巨大的研究对象,而这是一项花费巨大且容易犯错的任务。降低K值门槛也会给整个网络节点的数量带来指数增长而非线性增长。虽然目前K核心法的实际运用非常少,但是我们相信,在界定网络边界的实证研究中,它具有巨大的潜在价值。最近网络分析软件的发展应该便利了K核心法的使用。

基于事件的策略

这一方法在划定网络边界时,根据发生在特定事件和地点中的行为参与者而定(Marsden, 2005; Scott, 1991:60)。例如,弗里曼和韦伯斯特(Freeman & Webster, 1994)在南加州海滩观察休闲人士时,将范围限定在一个月内至少有三天会来海滩的人士。他们共收集了353个事件的参与人,其中包括43位规律性光顾海滩的爱好者。基于事件参与的涵盖

规则完全不同于位置策略法和关系策略法，后两者都忽略了行为人的行为。

由于基于事件的策略依赖具体的行为或事件来确定网络边界，因此如何对这类行为进行合理归类就显得非常关键。对于基于事件的分析来说，最让人畏惧的便是为选择能够回答具体研究问题的事件提供充分的理由。研究人员能够涵盖那些中立观察者认为是重要的，或者知情参与者认为是重要的事件。基于事件的边界界定方法易于因忽略某些重要的行为和参与人而导致获得不完整数据或者缺失数据。对于依靠一个单独事件来界定网络边界的研究者来说，这个问题尤其敏感，因为许多重要的行为人也许根本就没有参与该事件。因此，对多个事件进行观察，一般来说也许可以获得一个更加全面的分析网络。由于每一个事件都包含着一个潜在的独特的网络小团体，因此同时考察多个事件便得到一个可延展的网络边界：每一个事件都产生一个独立的网络，而其中的参与者部分地与其他事件的参与者相重叠。将所有事件的参与者集合起来，便得到一个更加广泛的网络，这一网络能够更好地回应研究者的问题。例如，为研究大学宿舍里的学生互动情况，弗里曼和韦伯斯特(1994)对咖啡馆和宿舍这两个不同环境中的社交聚会的事件参与者进行了观察。

第2节 | 数据收集过程

这一节讨论构建数据库的几个基本的网络数据收集过程,包括单个和多个提名生成法,全部个人网络关系的测量,位置选择法和资源选择法,以及档案资料的使用。

单个和多个提名生成法

提名生成法常用于自我中心网络关系研究,其通过调查问卷从每一个主体受访者那里收集与主体有直接联系的关系客体间的相关关系的信息(Marsdenm, 1987)。自我中心网络关系研究者依赖于两种调查工具:提名生成法和姓名阐释法。提名生成法确定调查对象的关系客体,姓名阐释法则获得每一个关系客体自身、关系客体之间的关系,以及其与关系主体之间的关系的信息(Marsdenm, 2005)。首先,主体受访者被要求提供一定数量的人名,这些人名与其有某种特别联系。接着受访者会被问及他/她所提供的人员之间是否有特别联系。最后,受访者还得提供每一个关系客体的具体特征,比如年龄、性别、种族和教育程度等。

根据研究目标的不同,自我中心网络研究可选择利用单个或多个提名生成法。单个提名生成法依靠一个调查问题

来引出关系客体的名字。1985 年美国综合社会调查中关于美国人的核心讨论圈的问题模块，生动地体现了这一方法设计（Burt，1985；Marsden，1987）。在 GSS 问卷中，为了引出受访者的知心朋友网络，访谈员会问：

> 大多数人总是会与其他人讨论一些重要的事情。那么，请你回顾一下过去六个月当中，你与谁讨论过重要的事情？你只需要说出他们的名字或者名字缩写就可以。

访谈员记录每个受访者说出的六个名字，然后继续问到："你感觉跟所有这些人的关系都是一样的吗？（如果不）那么其中哪一个是你觉得关系特别近的？（可以接着试探再问：还有其他人吗？）"之后，他们会问受访者关系客体之间的联系：

> 请想想你刚才提到的那些人之间的关系。一些人之间也许完全陌生，他们即使在大街上碰在一起，互相也认不出来。一些人之间关系也许特别亲密，跟他/她和你的关系一样好甚至更好。那么，现在想想甲和乙。他们之间完全陌生吗？或者他们关系特别亲密吗？（试探问：跟他们和你的关系一样亲密？甚或更亲密？）

受访者还要求回答：他们与关系客体认识多长时间了？平均间隔多长时间交谈一次？与受访者之间是什么关系（比如配偶、父母、子女、邻居、同事、朋友、顾问）？同时他们也要

报告每一个客体的性别、种族、教育背景、年龄、宗教和政党认同。

1985 年 GSS 的受访者平均只有 2.94 个可与之讨论重要事情的人，而这些核心讨论圈成员的 55%是家庭亲属(Marsden，1987)。自我中心网络密度测量均值为 0.61，意味着多数关系客体相互之间认识(自我中心网络密度测量关系主体所感知到的关系客体间的关系强度，其中 0 值等于完全陌生，1 值等于非常紧密，0.5 值等于其他)。美国人的核心讨论圈在年龄和教育背景方面比一般人群更具有同质性。性别多样性均值为 0.68，意味着在绝大多数核心讨论圈中男女性别是比较平衡的。相反的是，种族/民族异质性均值为 0.05，这一低值意味着绝大多数关系客体都是同一种族的，马斯登(1987)将这一现象归因于大部分的亲属间提名。

2004 年美国综合社会普查沿用了 1985 年 GSS 的提名生成法，但是网络规模均值从 1985 年的 2.94 下降到 2004 年的 2.08，这是一个主要的变化(Mcpherson，Smith-Lovin & Brashears，2006)。1985 年，一个受访者平均有 3 个知心密友，但 20 年之后这个数字却变成了 0：几乎近 1/4 的受访者无人可与之交流重要事情。非亲属关系连接点的巨大下降导致社会网络更加集中在配偶与父母中，而与公益组织和邻居的联系越来越少。受访者与关系客体的同质性持续偏高，伴随着教育背景异质性的下降，种族异质性上升。分析推测，美国人口特征的变化某种程度上可以解释美国人与日俱增的社会孤独感。

阮、弗里曼、戴、潘和张(Ruan，Freeman，Dai，Pan & Zhang，1997)利用 GSS 中的单个提名生成法于 1986 年和

1993年在中国天津进行了两次调查。在这两次调查中,天津受访者普遍提出了更多的核心讨论圈成员(均值分别为4.58和3.30),这比美国1985年的均值2.94高出许多,而且天津人的核心讨论圈成员中亲属要更少一些。相对于1986年,1993年的天津受访者核心讨论圈中较少提及同事和亲属,但是更多提到朋友。这些变化反映了中国自1978年以来社会宏观结构的变迁,其间中国利用更加灵活的市场雇佣制不断替代终身雇佣制。伴随着变迁的持续,人们逐渐发展了更多工作圈之外的社会关系,因此导致核心交往圈中其他关系人的加入机会的增加(Ruan et al., 1997)。

1985年GSS调查问卷在利用提名生成法时并没有关于所谓"讨论重要事情"的内容提示,而是将解释这一术语的任务交给了受访者。这种模糊处理可能使不同受访者对此有不同解释,从而影响自我中心网络的结构或组成(Bailey & Marsden, 1999)。例如,受访者如果将"重要事情"界定为个人和家庭事务,那么他将更多地提名联系紧密、有亲属关系的网络成员,而将"重要事情"界定为与工作或政治有关事项的受访者则较少提名亲属成员。为了仔细研究受访者在解释"讨论重要事情"时的认知过程,贝利和马斯登(1999)使用共时有声思维方法,对50个受访人进行调查。首先问这些受访者提名生成法使用的标准问题,接着立即用各种试探方法研究他们的思维过程。大多数受访者将之解释成个人问题,如家庭问题或个人间的问题。虽然先前的调查问题会明显引导受访者用其他方式解释"重要问题",但是这些不同的解释并不会形成实质性不同的网络构成。为了方便将来运用单个提名生成法,贝利和马斯登(1999)提出了四种可

选策略。第一种策略将定义内容与诱导关系客体分开:受访者首先被要求根据自己的理解定义重要事情,然后根据其定义来提名客体。第二种策略使用举例法来解释重要事情:研究人员提供一些例子,来帮助受访者定义有关重要事情。第三种策略中,研究人员为每一位受访者明确解释重要事情的含义。第四种策略则利用调查问卷问题次序的不同排列,来弱化有关问题背景对受访当事人理解重要事项的影响。

虽然单个提名生成法产生一个关系客体的核心集合,但是它所得到的联系人只是利用多个提名生成法所得名单的一部分。研究人员通常会对单个提名生成法所得的核心问题之外的各种常规行为感兴趣。在一个关于"朋友"概念的有趣研究中,费希尔(Fischer, 1982)发现人们使用这一相同的概念来描绘完全不同的关系。费希尔没有采用单个提名生成法,而是采用了多个提名生成法。这一方法包含九个事项,如家务、询问合适的贷款数额、社交、讨论有关工作、业余爱好以及其他个人事情等等。在北加州的调查中,他从每一个受访者处获得数量不等的客体名单(从 2 个到 65 个),平均下来每个受访者提供了 18.5 个名字。而且,受访者将其中 11 个人视为朋友,很显然各人对于"朋友"的定义五花八门,包含了各种人际关系。在以色列进行的同样两项调查,平均得到 14 个和 11 个姓名(Fischer & Shavit, 1995)。结果发现,以色列人比北加州人有着更高密度的自我中心社会网络。

为了比较单个和多个提名生成法,阮(1998)在中国天津调查时使用了 11 个提名生成法。除了利用 GSS 中的重要问题项外,她还利用了一些工具性连接点,比如家务和借贷,还

有一些表达性关系，比如社交、谈心和寻求建议。GSS提名法平均得到3.30个天津人的关系客体姓名，而其余十种提名法总共得到8.17个名字。而且，受访的中国人将“讨论重要事情”当做社交性表达事项，大多数受访者将与其社交的朋友（出去聚餐、购物或者游玩）视为他们讨论圈中的成员。相反，由工具性提名生成法所得到的名单更少被涵盖在核心讨论圈网络中。

由“认识”提名生成法可得到最大的自我中心网络。这种提名法要求受访者报告其认识的所有对象。奇尔沃斯、约翰逊、伯纳德、雪莱和麦卡蒂（Killworth, Johnsen, Bernard, Shelley & McCarty, 1990）的研究表明，在美国和墨西哥，这种认识提名法可以产生1500名认识对象。然而，为了获得与受访者有实质性联系人员的可操作名单，必须对此施加一些严格条件。坎贝尔和李（Campbell & Lee, 1991）总结了四种限制提名生成法的典型条件：(1) 角色/内容限制，将受访者限定在一种或几种关系中进行提名；(2) 地域限制，将受访者所提关系人限定为居住于一定区域；(3) 时间限制，要求受访者所提关系名单限定在一段可回溯的时间段内；(4) 数量限制，让受访者在符合标准的人员名单中只提出N个关系人（比如，三个你最好的朋友）。大多数提名生成法综合采用上述四种限制条件的其中几种。例如，坎贝尔和李（1991）请690位纳什维尔受访者利用其所居住位置地图，说出离他/她所知道的最近的9所或10所房子邻居的姓名。之后他们又要求受访者确认在过去的6个月中与之聊过至少10分钟或到访过该家的邻居。相比其他几种提名生成法，1985年GSS“讨论重要事情”提名生成法在内容、数量和时间段上

的限制最为严格,因此也只得到一个网络密度最高的最小自我中心结构网络规模(均值3.01)。费希尔的北加州调查,利用多个提名生成法,获得了一个最大规模的网络(均值为18.5),但是其网络密度也最小。纳什维尔调查中的关系持续时间是最长的,因为没有设置空间方面的限制,GSS调查中的关系联系频率是最高的。坎贝尔和李(1991)的研究使人们了解到,通过提名生成法所施加的限制条件是如何塑造一个目标网络的结构。

位置选择法和资源选择法

社会资源理论的基本假设是,社会结构——由财富、权力和地位来定义——是金字塔型和分等级的,社会资源及其取得路径正是嵌在这种社会结构之中(Lin, 1982; Lin & Dumin, 1986)。虽然有许多研究考察了社会关系连接点是如何成功地催化功利性行为的,但是一直以来却很少有研究者对取得社会资源的路径进行分析。为了解决如何衡量行为人的社会资源这一实证问题,位置选择法要求受访者报告其是否在某一社会位置上与他人具有联系。处于社会等级结构中的某个位置是社会资源的一种合理指标,在此意义上调查某个位置上的人的个人交往关系不仅可以显示其拥有何种社会资源,还可以显示其获得社会资源的途径。

如何通过位置选择法反映社会资源获取途径,主要依赖社会位置的选择。林和杜敏(Lin & Dumin, 1986)从1970年美国普查职业分类指标中选择了20种频次最高的职业。这些职业在白领和蓝领工作领域的上层和下层均有所分布。

受访者报告其是否在这些职业位置上有任何社会关系，比如亲戚和朋友（指有较紧密联系的人）或者认识的人（指有较弱关系的人）。利用1975年数据库（含位置选择法）来考察寻找工作的过程，林和杜敏（1986）发现，关系力度与较好职业的获得这二者之间的联系与受访者的出身是相关的，比如受到父亲职业的影响。当受访者出身社会上层时，强的和弱的社会关系连接点对于获得较好职业的作用基本平等。但是当受访者出身地位低时，弱社会连接点比强社会连接点能提供更高社会地位的工作机会。随后，林、傅和宋（Lin，Fu & Hsung，2001）将位置选择法用于研究台湾的职业声望与收入的关系。他们的报告指出，台湾社会在获取社会资本方面呈现出性别不平等现象。社会资本更有利于男性获得声誉好的职业和高的收入，然而对于女性来说，她们更依赖人力资本（比如教育）来获取较好的工作和收入。

哪些职业要涵盖在研究范围之中，这是利用位置选择法进行研究的关键点，当然这得根据研究问题来定。不同社会阶层的文化差异与社会网络关系是如何联系的？埃里克森（Erickson，1996）利用位置选择法对这一问题进行了研究。她抽取了三个主要的社会阶层的维度，即对财产、组织和技能的控制。由于她的受访者全部都在安全行业工作，因此基于技能控制的19种职业均反映了他们安全行业的特性（8个专业工人，4个蓝领工人，4个警察，2个商业管理者，1个企业主）。访谈员给受访者一些提示："现在我要问你，你是否认识多伦多地区某一行业里的人？比如，你是否认识任何律师？请列举一些你知道能与他们聊得来的人员，即使你与他们并不亲近。"假如受访者回答说他们认识某个职业里的人，

访谈人就会问他们关系的亲密度，比如是否“仅仅认识”，是“作为朋友关系的那种认识”还是“作为亲属关系的那种认识”。在一项基于性别的社会资本的研究中，埃里克森(2004)根据加拿大普查的数据，特地挑选了 15 种男性主导或女性主导的职业。她问受访人是否认识每种职业中的女性或男性。埃里克森(2004)发现，男性认识由女性主导职业中的人的可能性比女性认识由男性主导职业的人的可能性更大。而且，因为男性在许多社会部门都占据较为重要的职位，所以男性在社会网络中的优势难以改变，尽管相对于女性，两种性别都更倾向于与男性有更加多样的关系联结点。

位置选择法得到的自我中心结构网络，能用于测量个人在社会等级中与多个职业的连接关系。社会资源依照职业结构的金字塔型和分等级的不同层级进行分配，在此意义上位置选择法所得到的不仅仅是人们获得不同职业的途径，还是人们获得不同社会资源的途径。然而，人们常常获得超层级的关系客体提供的功利性和表达性帮助，而位置选择法却将选择限制在某些职业之中。研究人员常常对个人的社会资本宽泛地定义为人们从各种社会联系中获取的所有类型的帮助。资源选择法主要关注个人社会资本，由函数表达为 $SC = \sum_j S_j$，其中 j 指社会资源项，S_j 指测算这类资源的可获得性(Van der Gaag & Snijders, 2004)。与位置选择法询问当事人是否与特定职业有联系不同，资源选择法询问人们是否认识可利用某种具体资源的人。范德加格和斯尼德斯(Van der Gaag & Snijders, 2004)询问荷兰的受访者，“请问您是否认识谁可以帮助做……”(有 35 种选项，从“修自行车”到“礼节性拜访”)他们报告说，17 个资源选择项目形成了

四个亚层次，这些亚层次具有内在的同质性，但可关联性较弱。其中测量获得教区牧师这样有声望的职位的亚层次，与林和杜敏(1986)的位置选择法非常相关。但是，获取信息和获得工具性支持比如帮助搬家)的两个亚层次与位置选择法相关性很弱，因此形成了独特的社会资本维度。

测量全部个人网络

一些研究人员对测量全部个人网络感兴趣，因而创造出多种提名生成法来达到目的。全部个人网络指的是行为主体认识的所有关系。依据具体方法，全部个人网络规模可达250个至5000个(Freeman & Thompson，1989；Killworth，Bernard & McCarty，1984)。

在名单核对法中，研究人员首先随机选出几个姓名(或姓或名)，然后将这些姓名读给随机挑选出来的受访者，问他们:“你认识叫这个名字的人么?”当受访者说认识某一个名字时，访谈者便问他/她有关该人名的信息。麦卡蒂、伯纳德、奇尔沃斯、雪莱和约翰逊(McCarty，Bernard，Killworth，Shelley & Johnson，1997)利用一组50个名字的名册访谈了793位佛罗里达民众，采用的就是这一方法设计。访谈者请受访者报告与名册里名字相同的朋友的联系。研究要求受访者通过外貌或名字知道对方，并且他们在最近两年里有过联系。麦卡蒂等人(1997)通过这种办法获得了最多14个关系客体的一个名单，但发现这个名单样本中黑人、西班牙裔和亚裔的代表性偏低。然而，他们认为在发掘一个更具代表性的全部个人关系网络样本上，这一方法依然优于其他办

法。认识名单所列某一名字的受访者比例,与美国人中拥有同一名字的人口比例,二者高度相关,这正表明该方法的偏差是不严重的(McCarty et al. , 1997: 311)。

在反向小世界方法(Reverse Small-World Method, RSW)中,研究者首先创造一个虚拟的目标,其中包括名字拟造的人员、随机分派的地址、年龄、性别、爱好、成员资格、职业以及其他社会经济特征。之后他们要求知情人提供有关的关系客体,而这些人员能直接提供一个信息给虚拟目标,或者形成一个指向目标的网络连接点。明确了关系客体之后,研究者便向知情人士也就是受访者询问这些关系客体的个体特征、与知情人的关系、关系客体之间的认识程度。奇尔沃德等人(1984)利用 RSW 方法访谈了佛罗里达杰克森维尔的 15 个知情人士,获得了 100 个虚拟美国人和 400 个虚拟国际人士的信息。这些知情人士提供了大约 134 个关系客体,通过这些客体与这 500 个虚拟对象联系起来。大多数关系客体(86%)是知情人士的朋友,而且超过一半都是男性。

弗里曼和汤普森(Freeman & Thompson, 1989)认为,RSW 方法只能获得一个人全部关系网络的一部分,因为知情人不可能提出超越虚拟目标数的其他关系客体。为了解决这个问题,奇尔沃斯等人(1990)将 RSW、名单核对法和 GSS 提名生成法综合起来。他们发现,佛罗里达人的个人全部网络关系规模平均是 1700 人,而墨西哥城居民的平均值是 600 人。利用同样的数据资源,伯纳德等人(1990)调查了上述不同方法中的重合部分。他们发现 GSS 提名生成法和多重社会资源提名法总共才获得 RSW 提名法所得个人全部网络关系数量的 18%。名单核对法(通过姓氏核对)得到的

个人网络关系数量最大。

档案资料

档案资料数据为社会网络分析提供了许多有用的信息，虽然这些信息原先并不是为了网络研究而收集的。与专门调查相比较，档案资料相对而言使用成本低，对受访当事人不会造成时间和精力上的负担，而且当数据是重复收集的话，档案资料有可能保留一些非常难得的纵向信息。档案资料来源形式多样，包括网页、个人信件、专利摘要、书籍和论文文献，以及电脑网络交流文本等。几个在线数据供应商，比如 Lexis/Nexis、Free Edgar、SDC Platium 以及美国专利商标局等，存储了大量有关公司董事会、战略合作伙伴、小企业主以及专利摘报等信息，极大地方便了研究人员对商业组织网络的研究。

档案资料检索新技术的出现总是层出不穷。利用互联网这一巨大丰富的个人信息库，亚当米克和阿达尔（Adamic & Adar, 2003）发展了一种技术，用来发掘其中的人际关系网络。他们用一种客户端软件，调查两个大学（斯坦福和麻省理工）服务器上个人网站的网页内容。他们关注每一个主页的文本、外部链接、内部链接和邮件名单。他们这种发掘互联网内容的创新尝试，不但了解到谁认识网站主人这类信息，而且还了解到他们的认识是在何种社会情境中发生的。例如，麻省理工的学生倾向于通过大学生联谊会及其下属组织来互相认识，其中提名最多的十个社交场合有五个是大学生联系会之类的组织。比较而言，斯坦福的学生之间的认识

渠道更加多样化，比如通过研究团体、宗教组织、文化交流机构以及学生联谊会等。

个人书信和电子邮件带有大量有关个人交往的信息，因此分析检查这些资讯文本会得到有关个人网络和社会结构方面的丰富信息。这种方法对历史学家尤其有利，因为他们所研究的古代社会现在是不可能去调查的。例如，亚历山大和达诺夫斯基(Alexander & Danowski，1990)通过西赛罗(Cicero)的书信，对古罗马的社会结构进行研究。西赛罗是古罗马著名的演说家和有影响力的政治家，他跨越两大社会精英阶层：无官职的"骑士阶层"和有官职的"元老院阶层"。研究者查阅了280封西赛罗18年间与他的熟人、朋友、亲属的通信。他们的数据管理程序记录了524个人之间的1914种关系，包括关系双方的姓名和官职，以及诸如送礼、预约、中间调解、帮助、负面描写、拜访等关系内容。该研究一个主要的发现就是，虽然元老院成员和骑士在某些事情上互相对立，但是他们似乎构成了一个单一的、整体融合又互相牵制的社会阶层。

期刊和专利都含有大量引文名单，因而为社会网络研究提供了另一个很好的数据来源。哈根斯(Hargens，2000)对三门自然学科、三门社会学科、一门人文学科的引文模式进行了研究。对每一门学科，哈根斯首先确定一个期刊清单，这一系列期刊使用引文索引服务和专家咨询服务。然后他记录每一篇论文中引述其他文章的参考文献或脚注，从而为每一个学科领域建立了一个全面的引用关系网络。哈根斯发现研究者因各种目的而引用他人文献，从认同一个普遍的理论流派到纠正某一具体的计算过程。特别需要指出的是，

自然科学的文献引用大多数是为了拒绝或更正某一具体实验方法，而社会科学家绝大多数是为了认同或挑战某一学科的基本假定。在对之前的文献进行争论方面，人文批评领域往往要比自然科学和社会科学界更为盛行。

对于有志于研究知识流动的人来说，专利授予过程产生了可资研究的丰富数据。在对专业技术市场的研究中，波多尼和斯图尔特(Podolny & Stuart, 1995)深入分析了专利摘要的文献引用。为了获得专利，技术发明者要向美国专利局提交申请书。作为申请程序的一部分，他们必须列出作为此次申请重要基础的所有已授权专利。专利审查员需要确认引文清单包括了所有之前的相关发明。利用 Lexis/Nexis 在线数据库，波多尼和斯图尔特(1995)分析了自 1976 年至 1991 年间授予全球半导体技术创新者和制造者的所有美国半导体设计专利。他们将专业技术市场看做由一系列技术创新构成的，这些技术创新之间是由一种创新到另一种创新的知识流动链进行连接的。因此，每一项授权专利都是一个进入已有技术市场的新的重要的技术创新。一项创新技术的专利摘要文献，提供了与之相关的所有前人的基础性创新信息，而基于此项发明的后一项专利也表明它们之间的联系是建立在此项重要创新之上的。因此，一个专业技术生态位构成了一个自我中心网络，该网络是由某个重要创新及它与其他专利发明之间的联系形成的。

虽然调查文献引用清单可以构筑学者或专利人之间的关系网络，但是其中有两点重大差异(Meyer, 2000)。第一，期刊论文倾向于引用大范围的已发表文献，而专利引用范围较小，往往只引用对此项创新有重要贡献的前人发明。第

二,期刊论文引述目的较广,比如赞赏有关论著、纠正前人工作或者与某人争论某个观点(Meyer, 2000)。相反,专利摘要文献引述的目的就单纯多了,就是为了感谢前人的工作为当前的创新提供基础。因此,利用专利文献和期刊论文所建构的社会网络可能在内容上有根本性差别,需要研究者仔细甄别。

政府机构和经济组织在雇员间的交往以及有影响力的组织间的互动上,已经累积了大量的信息。然而,这些有价值的数据资源普遍被忽视,组织理论者对此也少有问津(Perrow, 1986:172—173)。一个例外是贝克(Baker)和福克纳(Faulkner)的研究,他们利用了基福弗(Kefauver)特别委员会报告。这一报告是该委员会在美国参议院反垄断司法委员会上关于商业合谋的证词。通过这个报告,作者揭示了非法商业关系网络的结构及其动因。

自20世纪90年代以来,随着电脑网络技术的迅速应用,电子化交流不再局限于少数技术专家,而是迅即扩展到普通民众。新技术用于控制或激励网络使用者的交流,从而让研究人员便于研究它、揭示它的模式。社会网络研究者通过检视这些电子化交流(CMC)系统,比如电子邮件、电子公告栏、语音信息和其他交友软件,获得了丰富的社会网络数据(Rice, 1990、1994; Rice, Borgman, Bednarski & Hart, 1989)。由于监控程序普遍都可存储所有的CMC系统的交流,因此它的网络数据便构成了此系统使用者间所有交往的一个普查。然而,对于网络分析人员来说,检查分析每一项交往,在计算时间和成本上将会是不可承受的。更为重要的是,由于这种电子化交流在内容、时间和目的上都呈现多种

形式，因此并不是所有的交流数据都是有用和相关的。CMC研究者通过三个原则甄别数据：(1)在基于角色的研究中，研究者过滤那些基于角色——比如经理或者会计——的信息发送者或者接收者；(2)在基于时间的研究中，研究者设定一个时间段，在此时间段内收集所有交流信息；(3)在基于内容的研究中，研究者过滤一些内容和主题，而只选择那些符合研究问题的信息。客户端计算机软件通过关键词对大型网络数据库进行快速搜索，极大方便了基于内容的研究。

与自填式问卷调查相比，CMC分析方法的强制性较弱，且具有更小的测量误差。例如，当两个人自填的关系不一致时，要发现这些测算错误的来源，比如回忆不准确、回答有偏、问卷用词问题等，就变得几乎不可能了(Rice，1994：176)。然而，计算机控制网络数据能够精确地反映CMC信息的内容，因此有关相互关系的假定就能够得到明确的区分和验证。虽然CMC数据比自填式问卷得来的数据更精确，但他们并不一定更有效。他们通常展现人类交流的不同方面(Rice et al.，1989)，例如，电子邮件交流能更准确反映工作行为而非朋友间和家庭成员间的社交行为。作为使用工作服务器电子邮件系统协议的一部分，雇员普遍都签署了协议同意自愿分享他们的交流。因此，依据研究主题的不同，研究者应当对CMC系统产生的数据保持警醒。对某一公司服务器用户的电子邮件交流进行监控检查，可以获得关于工作联系结构的丰富数据，但是却很难获得雇员业余时间的关系结构数据。单独依赖CMC系统也可能会错失一些非常重要的细微交流之处。大量信息内容是不重要的例行公事，偶尔会有一条非常重要的信息，忽略这样的信息也许会产生严

重后果。因此,研究者应当对CMC数据分析保持警惕,因为这种方法虽然揭示了有价值的背景线索,但同时也丢掉了一些交流行为中的重要的东西(Rice, 1994:177)。

CMC技术的最新发展大大延伸了网络的运用,便利了社会网络分析(Berman & Brady, 2005)。网络基础设施包括诸多网络工具:比如高性能计算机和网络,个人电子设备(即时信息工具、黑莓手机和个人数字化助手),数据获取、管理和分析服务,可视化设施等。网络基础设施为社会网络分析提供了大量新的机会。例如,高性能计算机服务器记录万维网和个人电子设备使用者的文本信息。对这些文本信息进行分析会得到重要的社会交往模式和网络构造。大量的机会同时也带来巨大的挑战,从如何获取、存储、分析海量数据这些技术问题,到如何确保获得知情同意权以及保护隐私等法律和社会的问题。因此,网络基础设施所支持的社会网络分析,要求多学科之间进行前所未有的合作,包括社会科学、法学、电子学和计算机科学等。

亚当米克和阿达尔(2003)利用网络工具获取和分析巨量数据的研究就是一个很好的范例。这些数据信息在网络空间流转,通过对它们的分析,深化了我们对社会网络中某些重要事件的理解。为了研究小世界现象,他们从惠普公司实验室员工的电子邮件和超过2000名斯坦福大学学生的社交网页中,收集了大量信息。这个交流网络包含流通在430个人之间的4000封电子邮件。这些人在3个月的观察期内,每人至少交换过6封电子邮件。网页网络数据来源于一个社群网址Club Nexus,它邀请斯坦福大学学生加入该俱乐部并且登记他们的朋友圈。登记过程中也要求登记人报告

他们的专业、性别、个人特征、宿舍地址、爱好等。朋友也被请求回应登记在Club Nexus的关系。未得到回应的关系将从数据库中删除。为了定位交流目标，亚当米克和阿达尔对被观察电子邮件网络使用了三种搜索策略，每一种都反映了一个潜在的理论观点。高相关度寻找法选择那些通过认识很多人来了解研究对象的个人。第二种便是将信息传递给在组织等级架构层级中离研究对象最近的联系人。第三种便是将信息传递给在空间距离上离研究对象最近的联系人。在两种网络中模拟这三种策略后，亚当米克和阿达尔认为第二种——位置搜索发现了实验室网络中关系双方的最短路径。然而，没有任何一种搜索方法得到了俱乐部网络的最短的合理路径。

第3节 | 认知社会结构

虽然关于社会网络最普通的问题之一可能是:“谁认识谁?”但是,与认知社会结构(Cognitive Social Structure, CSS)分析最相关的问题也许是:“谁知道谁认识谁?”研究者使用CSS数据调查知情人对于网络关系的认知的变化。虽然知情人可能是外在的观察者,但是他们也是其所报告的网络的参与者。为了创建一个CSS数据库,研究者要求每一位参与者报告他/她对于每一个对偶网络关系的主观认识。例如,使用一个涵盖所有双方行为参与人的名单,去核对所指关系的感知性存在。对于某一个具有N个行为人的网络,一类关系的完整CSS数据库含有N个认知图,这些认知图标明了每一个参与者—知情者视角下关系连接点的显现与缺失。鉴于可能需要大量的时间和回忆的负担——每一个知情人需对具体的有向对偶关系进行 $N^2 - N$ 次判断,CSS数据收集工作常常限于非常小规模的网络,或者一个有限的任务范围(例如,“谁是X的三个最好的朋友?”“Y最信任哪五个人?”)。举例来讲,克雷克哈德(Krackhardt, 1987)收集了21个管理雇员中的两种类型关系的CSS数据——建议寻求和友谊,得到了两套各21个认知图,每套展现了 $(21)^2 - 21 = 420$ 个有向联系。每个自我中心网络都是一个CSS特例,其

中知情人只描述所观察到的他/她的直接联系人之间的关系，而不是与所有其他网络成员间的联系。

CSS数据库的两个普遍运用是:(1) 判断参与者对网络结构认知的一致性程度;(2) 通过与某些准则进行比较，评估所观察网络结构的认知偏差。一致性是指两个或更多的知情人对社会关系判断或评估的同意程度。一个重要的理论问题是，具有相似个人特性(比如种族、性别或年龄)的行为人之间的一致性程度是否更高? 或者共享关系连接点(或位于相似网络位置)的行为人，是否更有可能在网络关系的认识上达成一致? 偏差指知情人通过某一准则的比较，达到对网络认知的准确性程度(更多关于知情人偏差的问题，请参看后面章节)。其中一个准则是所有其他参与者所认知的网络的集合。另一个准则是从参与人行为的直接观察中所获得的关系数据。对自填式问卷测量与关系行为测量之间的较大差别，网络分析者已经观察很长一段时间了，并且已提出解决这种差别的办法。虽然大部分网络数据是通过面对面或者电话访谈获得，但是大部分研究人员很少注意受访人或知情人的认知社会结构，而这种认知社会结构会塑造知情人的回答。分析者一般将所有的反映当做明确客观的社会事实。然而，一些学者已经寻求理解和测算基于受访者认知获得的信息与行为观察之间的差异来源(如 Batchelder, 2002; Carley & Krackhardt, 1996; Casciaro, 1998; Freeman, 1992; Johnson & Orbach, 2002; Koehly & Pattison, 2005; Krackhardt, 1987)。

认知社会结构也帮助研究者确认受访者自我填答时出现的系统性认知误差。昆巴萨尔、罗姆尼和巴切尔德(Kum-

basar, Romney & Batchelder, 1994)访谈了25个计算机硬件工程师,构建了一个认知社会结构。每一个知情人都被提供给一个随机抽选的工程师的名字,然后被要求从整个25个工程师名单中标注一个人,他们认为这个人是随机所选工程师的朋友。这一过程对所有的名字重复执行,包括知情人自己的名字。然后利用代数矩阵方法(参见第4章),获得的25个认知图能够整合到一个横跨所有工程师—知情人的全幅网络。昆巴萨尔等人(1994)将自我认为的朋友网络与全幅网络相比较,得出结论认为,个人倾向于认为更亲近他们自己网络表征的中心,而不是全幅网络的中心。另外,个人倾向于构建一个具有认知连贯性的网络,也就是说,他们将自己看做由朋友所包围的中心,而这些朋友也有大量的朋友圈子。

知情人倾向于构建一个具有认知连贯性的网络形象,这一发现并不新奇。关于认知心理学的大量文献认为,当人们遭遇关系不平衡时,他们会经历心理或情绪上的压力(Heider, 1958; Newcomb, 1961)。随后,面临不平衡关系的个人开始启动"平衡计划":他们改变他们的认知,以求获得这一关系的平衡性认知图像。然而,这种认知的调节主要发生在朋友关系间,尤其是亲密朋友关系间,而在非朋友关系间人们似乎不受这种认知的不平衡所影响。实际上,当个人之间没有直接联系时,认知的不平衡很少导致心理上的不平衡。因此,知情人虽然尽力保持他们的心理平衡,但是他们还是更易于在朋友间提供扭曲的和有偏差的信息,相反他们对于疏远关系的报告更少受到偏差和扭曲方面的质疑。

与这种传统认识形成对比的是,克雷克哈德和凯尔德夫

(Krackhardt & Kilduff，1999)从泰勒和费斯克(Taylor & Fiske，1978)的认知吝啬者模型中得到启示，认为知情者一般都缺乏与他们并不直接联系的远距离关系人的了解。通过动员平衡计划来填补信息空缺，知情者在报告远距离关系和亲密关系时都会存在偏差。但是，两种偏差的动因是不同的：对于亲密关系来说，平衡计划是用于获取心理和情绪上的平衡，但是对于远距离关系来说，平衡计划是用于弥补有关联系信息的缺失。克雷克哈德和凯尔德夫(1999)调查了四家公司的员工，请每一个知情人士报告其同事间的朋友关系。他们发现了网络结构与认知平衡间的曲线关系：员工认为他们的亲密朋友圈和最边缘关系圈要比那些中等关系的人更易于平衡。其中一个主要的启示便是，近距离关系和远距离关系都容易产生高度信息偏差，这是下一节的论题。

第4节 | 信息偏差

我们将信息偏差定义为自我报告的行为与真实行为之间的不一致。对于言词和行为之间的鸿沟,社会科学家早已开始进行观察。20世纪30年代早期,理查德·拉皮尔(Richard LaPierre)和一对年轻的中国夫妇开车横穿美国,途中他们在184家餐馆吃过饭,在66家旅馆住过宿,并且只被一家旅馆拒绝过。旅程结束后,拉皮尔给所有的餐馆和旅馆等发信,问他们是否"接待过中国人"。结果,92%的回复是"否"(LaPierre, 1934)。拉皮尔因此总结,偏见态度的行为表现是需要支撑条件的。在另一项研究中,克罗嫩伯格(Kronenfeld, 1972)询问刚离开餐厅的人,男侍者穿着什么样的服装?餐厅又播放什么音乐?许多受访者对服装和音乐提供了详尽的信息,但实际上这几家餐厅根本就没有侍者或没有播放过音乐。

信息偏差的来源多种多样。许多受访人对于准确报告他们自己行为的要求,并不具备处理这种大量信息的认知能力(Bernard, Killworth, Sailer & Kronenfeld, 1984)。另外还有两个方面的原因:知情人在认知上将某一种归类形式强加至某一不归类的附属模型上(Freeman, 1992);知情人矫正他们的认知以便维持亲密朋友间的平衡关系(Krackhardt & Kilduff,

1999)。一些研究人员认为,信息偏差并非随机的噪声,恰恰某些受访知情人在准确提供行为信息上的确要比其他一些人强。例如,一些研究表明,对受访主题的熟悉往往会有助于受访者对事件的回忆(Romney & Weller, 1984; Romney, Weller & Batchelder, 1986)。相反,另一项研究发现,虽然对论题知识渊博的受访者遗忘较少,但他们也容易同时创造出一些诸如并不存在的关系成员之类的错误(Freeman, Romney & Freeman, 1987)。还有分析发现,所有的受访知情人士都倾向于将自己视为更加靠近网络关系的中心,而并非如其他网络成员所报告的那样。因此,对某一网络组别层面的报告也许要比个人层面的报告来得更准确些(Kumbasar et al. , 1994)。最近几项关于自我中心网络的研究都对信息偏差问题进行了考量(如 Brewer, 2000; Feld & Carter, 2002; Johnson & Orbach, 2002; White & Watkins, 2000)。

对于将自填式问卷得到的数据作为网络关系人的真实行为数据进行研究的社会网络研究者来说,也许最引人注目的消息就是 H. 罗塞尔·伯纳德和他的同事们所进行的项目(Bernard & Killworth, 1977; Bernard, Killwoth & Sailer, 1981; Bernard et al. , 1984)。他们构建并比较了七套配对的交流网络数据库:由参与人自述建构的认知数据;由机器监控系统或"客观"观察者记录所建构的行为数据。这些数据库系统包括:两组华盛顿特区的聋人 TTY(文本电话系统,专为听力或语言不便者使用)使用者;弗吉尼亚、宾夕法尼亚和俄亥俄无线电协会的业余使用者;社会科学研究所的职员;西弗吉尼亚大学技术教育的一个研究生项目;大学学生联谊会驻地;一组电子信息交换系统(Electronic Information

Exchange System, EIES)使用者。例如,EIES 使用者记录自述数据(认知网络)和通过 EIES 电子监控系统所得数据(行为数据)。所有七组双方交流网络数据包括:"受访人与谁联系交流?""他们多长时间联系一次?""他们联系多长时间了?"通过对这些认知数据和行为数据进行比较,伯纳德等人(1981:17)发现大约一半的自我报告知情人在某些方面是存在问题的,并得出如下结论:

> 如果对其要求任何可接受的精确度,人们实际上并不能够说出他们在过去特定的一段时间内与谁交谈过……我们现在确信,关于交流的认知数据并不能够代表同等行为的数据。

其他一些学者认为,情况也并非像伯纳德和他的小组所描述的那样毫无希望。诺克和库克林斯基(1982:32)对"客观"观察者的准确性和中立性表示怀疑。他们提出如下一些问题:"在什么程度上这些中立的观察者打扰了正在进行的行为?""他能够多精确地记录双方的交流?""在非观察期他们之间的交流如何?"伯纳德等人(1984)后来调整了他们的例子,承认尽管在认知数据方面存在一些系统性扭曲,但通过一些改进因素可以减少信息偏差,比如让知情人更熟悉文化背景,或由访谈者给受访者提供相关回忆提示和帮助。他们建议"信息的不准确性不应该导致抱怨或绝望,而应该引至一个尚未开垦的研究新领域"(Bernard et al., 1984:513)。

信息偏差常常以其他系统形式发生。持偏差观点的受访者将自己视为更趋网络中心(Kumbasar et al., 1994),而

忘却边缘行为人，错误地回忆主要网络成员（Freeman et al.，1987），在认知上矫正他们的认识以达成他们与或近或远的朋友间网络关系的平衡（Krackhardt & Kilduff，1999）。信息偏差也往往由不准确的回顾所致（Bernard et al.，1981；Bernard et al.，1984）。不能准确回忆的知情者容易犯两类错误：(1) 忘记他们与谁交往过；(2) 错误地回忆实际上从未发生的事（Freeman et al.，1987）。而且，遗忘和错误回忆并不是随机发生的。知情人倾向于遗忘并不频繁的参与者，错误回忆频繁的参与者。弗里曼等人（1987）记录了参加 9 个座谈会的大学教师和学生。最后两个会议结束后 5 天，研究者调查了 17 位参加过这两次会议的知情人和 16 位没有参加这两次会议但是参加了至少两次之前会议的知情人。研究人员问他们参加和缺席会议的情况。这些知情人在 272 个选择机会（ $17\times16=272$ ）中犯了 141 个错误（115 个遗忘和 26 个错误回忆）。52%的错误率非常接近伯纳德等人（1981）的估算。伯纳德等人估计过半的知情人自我报告是不准确的。座谈会知情人在对其他参与人的了解上，差异较大，而这就影响到他们的错误率（Freeman et al.，1987：312）。与知之甚少的知情人（他们只在座谈会中与其他人交流过）相比，知识丰富的知情人（他们对其他人在许多方面都有所了解）更少遗忘，但更多地犯回顾性错误。因此，与通常的认识相反，知识丰富的人实际上更容易错误回忆。然而，这些知识丰富容易错误回忆的人却是会议的经常参与者。由于社会研究的主要关注点是如何测量各种重复行为的模式，因此就收集这些数据而言，这些知识丰富的人倒是很好的知情人（Freeman et al.，1987）。

知情人之间的高度一致常常意味着低误差和高信息效度。因此,罗姆尼等人(1986)提出一个共识模型,将知情人的能力(指知情人了解研究客体的程度)和知情的统一性、有效性联系起来。一个核心观点是:任何两个知情人答案之间的一致性,是每一个答案与真相相关联程度的函数。也就是说,相对于无效的回应,有效的回应更容易经常得到同样的答案。研究者要求知情人回应一系列问题,而对于这些问题只有研究者本人才有标准答案。他们用所答正确数除以答案总数的比例来评价每一个知情人答案的"真实"有效性水平,用共识模型来评价其一致性有效水平。他们发现知情人的真实性准确度与他们的一致性准确度高度相关,因此他们也认为一致性的答案往往等同于有效的答案。

总之,与伯纳德等人(1981)宣称的知情人自我报告总是错误地反映事实真相这一结论相反,对信息偏差的研究为如何更合理利用知情人提供了有益的建议。知识丰富的知情人能提供有关长期重复模式的无偏数据,同时他们也能够提供一致的答案,而这正意味着更高的有效性。个体认知相比于系统对所有网络行为人的综合认知而言,更偏向于将自己视为网络的中心。因此,系统认知要比个体认知更能准确把握一个人在网络中的定位。另外,自我往往会通过"平衡计划"对亲密朋友和较疏远的联系人产生认知偏差。那些利用自我中心关系法来收集网络数据的研究者,在分析亲密关系和较疏远关系时,应当采取一些矫正措施。

第5节 | 信度

从一般意义上来讲，信度是指特定工具用于重复测量同一事物时，每次都得到同样结果的一种程度。其中，有好几种测量信度的方法，包括观察者间信度、再测法信度、内在一致性信度（包含折半信度和 Cronbach α 信度）（Neuman，2000：178—185）。社会网络研究者常常使用再测法来判断知情人的信度。例如，对知情人提出或剔除讨论重要问题的联系人名册，再测法会在以后的某一时间再次就同一问题调查该知情人。比较前后两次反应的一致性程度，即可得知每一个知情人的信度。当两次调查中知情人都提供了完全相同的核心讨论圈名单时，即这两次测量能够完美对应时，就表示完全可信。相反，在一个完全不对应的测量中，知情人选择了与上次完全不同的名单，那就意味着完全不可信。在操作再测法时，一个关键的问题是两次测量之间的间隔。两次测量之间的不同答案，也许反映的是个人网络关系的真实变化。实际上，人们不停地发展新的人际关系，他们之间可以进行表达性的或实质性的帮助，比如讨论重要问题、搬家、照看小孩或者借钱等。方法论者认为，这种颠覆性问题可以通过缩短两次测量之间的间隔来减弱其影响（Brewer，2000）。间隔时间越短，个人网络产生的实际变化可能就越

小,因此这些变化对信度测量产生扰乱的可能性也就越小。

一个简单的但经常使用到的再测法信度测量是 Jaccard 系数,它指排除共同负面配对之后的同意部分。它是这样计算的:用知情人两次测量都提名的人数,除以两次测量中提名的总数[①](Brewer, 2000)。Jaccard 系数值从 0.00 至 1.00,表示零可信度和完全可信度。下面是一个含有 35 个姓名的虚拟名单,结果显示在交叉表中,单元频率按照 A—D 来表示。

	第二轮访谈	
第一轮访谈	提名人数	未提名人数
提名人数	A = 12	B = 5
未提名人数	C = 3	D = 15

在第一轮访谈中,知情人选择了 17 个提名;在第二轮中选择了 15 个提名;在这两轮中共有 12 个人被同时提名。因此,Jaccard 系数 = A/(A+B+C) = 12/(12+5+3) = 0.60。一个小组中的每一个知情人可能具有不同的信度,因此一个小组级别的 Jaccard 系数便是所有个人系数的均值或中位数。

知情人信度常常与有效性相关联(Romney & Weller, 1984)。假如知情人的自我认知报告只有一部分反映了他们的真实行为,那么知情人之间的一致性就往往能够表明知情人回忆的有效性水平。例如,针对网球比赛的问题,网球运动员应当比非网球运动员表现出更多的一致性。知识丰富的运动员也许能够完全一致地回答这些问题,而没有相关知识的运动员则会给出五花八门的答案。从这个中心观点出

① 两次都提名的人算一个人名。——译者注

发，罗姆尼和韦勒（Romney & Weller, 1984）再次分析了伯纳德等人的数据库，用于确认那些有效可信的“好的”知情人和那些无效不可信的“差的”知情人。他们发现，个人的信度与他/她的效度是正相关的。个人的信度通过个人回顾与组别整体回顾之间的相互关系来测量，效度则根据个人回顾与组别整体数据的一致程度来计算。另外，高信度的人比低信度的人更倾向于与其自我报告有更高的相关性。因此，主要的发现是，可以通过信度预测效度，具有高效度的知情人之间应该要比低信度的知情人之间，具有更多的相似反应。而且，知情人应当由他们的信度来加权：那些具有更高信度知情人的答案应该比那些低信度知情人的答案赋予更高的权重。仅就高信度人群这点来说，伯纳德等人的数据中的知情人提供了有效的自我报告，这从他们的回忆数据和综合观察互动数据的高度一致性能看出来（Romney & Weller, 1984：75）。

在一定程度上，知情者中忘却某些联系是不可避免的。因此，相对于比较自由回忆法，可以采取某些帮助回忆措施——比如提供线索或提示的名单——它们的可靠性问题更少。比较自由回忆法是指知情人在没有任何帮助的情景下进行提名。例如，琼斯和费希尔（Jones & Fischer, 1978）针对某种具体网络关系，比如“借钱”、“讨论工作有关的事务”，访谈了86个人。第一轮访谈中，受访者自由地回忆那些关系中的个人。第二轮，研究人员给每位受访者一个名单，并且重复了同样的问题。结果第二轮比第一轮多得到34%的人员名单，而且第一轮提名的27%在第二轮访谈中并没有被提及。

在另一项网络信度研究中,布鲁尔和韦伯斯特(Brewer & Webster, 1999)调查了217位住在一个大学宿舍楼中的人员。第一轮调查要求他们写下居住在这栋大楼的所有朋友的姓名。第二轮则为每个受访人提供了住在本楼的人员花名册,然后请受访人再次确认他们认识的朋友。利用花名册的提示,受访者平均多提出了20%的朋友名单。而先前忘记的名字,多是关系比较疏远的人,这也表明如果没有提示和线索,受访者一般会遗忘与他们关系相对疏远的对象。其中一个推论是,当一个研究人员将在两次独立的访谈中获得的自由回忆数据和基于帮助而得到的数据进行比较时,会出现结果非常不可信的情形。在这种情况下,低信度只意味着更高的效度:知情人在一个访谈中比在另一个访谈中得到更完整的提名。当然,信度并不保证效度。一个知情人在两次访谈中也许提名完全相同,因此获得完美的信度,然而每轮访谈都可能因同时忽略了同样的名字而产生了同样的偏差。不过,高效度的网络反应一般来说具有高信度。也就是说,假如一个受访者能够总是根据某一提名法则列举同样的朋友名单,那么这个受访人也很可能在不同的访谈时间里选出一致的朋友圈子名单。

由于社会网络数据的特有属性,尤其是在自我中心数据分析中,受访者的信度和效度测量与传统的其他种类数据的测量差异非常大(Marsden, 1993)。从较小且封闭的社会中抽取的自我中心样本中,知情主体常常提出许多与某一具体关系有关的相同的关系客体。知情人之间在选择相似关系客体上所体现的高度相关性,意味着知情人的高度可信。同样地,知情人对他/她的关系客体特征(如年龄、性别、教育背

景、社会地位等)的描述,与关系客体对这些特征的自我报告二者之间的高度相关,也意味着知情人的高度有效性(这个解释假定所有关系客体的特征都可从其自我报告中获知,从而用于作为比较知情人自我报告的事实数据)。

研究者对肯尼亚乡村妇女间的避孕话题网络进行过一个专门调查,揭示了这种自我中心网络方法的信度和效度问题(White & Watkins, 2000)。研究者在1994年12月访问了925位妇女,一年后再次访问了其中的743位。在两次访问中,这些妇女都被问及同样的问题:“你与谁商讨有关节育问题?”受访者随后报告其中四个随机抽选出来的关系客体的年龄、教育背景、收入状况以及节育措施的使用情况。研究人员也询问每一位关系客体自己的年龄、教育背景、收入状况和节育措施情况。怀特和沃特金斯(Watkins)发现,在受访者自我报告和关系客体自我报告之间,节育措施情况存在最明显的差异。相反,在一些可直接观察到的事项上则存在较高的一致性,比如“朋友家的房子是否有一个金属屋顶”或“她家里是否有套沙发”。而且,由受访者(提出了同样的关系客体)所提的关于关系客体的报告的相关度只有 $r=0.31$,只比受访者自我报告与关系客体自我报告的平均相关度(0.25)高一点点。这些发现表明,受访者对于关系客体节育措施的了解是既不可信也非有效的。第一轮由受访者提出的关系客体只有18%在第二轮中提到,这也表明了或者是低信度,或者是在受观察年份期间中关系网络的不稳定。

第6节 | 数据缺失

社会网络研究对数据缺失尤其敏感。在自我中心网络数据收集过程中,一个拥有 N 个关系客体的关系主体被要求报告 C_N^2 个非定向的对偶关系。具体来说,$C_N^2=\frac{N!}{2!\times(N-2)!}$,其中 $N!$(读做"N的阶乘")表示从1至 N 的所有正整数的乘积。例如,假如关系主体有5个关系客体,则

$$C_N^2=\frac{5!}{2!\times(5-2)!}=\frac{(1\times2\times3\times4\times5)}{(1\times2)\times(1\times2\times3)}=\frac{120}{12}=10;$$

因此,关系主体必须对这10种非定向关系给出有关信息。对于定向关系,关系客体间的连接点数量就是上面的两倍:$2C_N^2$;因此,在上述例子中,关系主体应当报告20种定向关系点。自我中心网络关系的关系应答率 R,等于所报告关系点的数量除以关系客体间所有可能的对偶关系数量。比如,关系主体报告了10种非定向关系中的8个,那么 $R=0.80$,或者80%;假如关系主体漏报了20种定向关系中的6个,那么 $R=0.70$,或者70%。

要计算一个完全社会网络的应答率是比较复杂的一件事。一个完全的社会网络由网络中 N 个行为人配对所组成的所有双边关系所构成。对于非定向关系来说,R 相对就简单些,因为当测量信度较高时,由双边关系中的一方来报告

便足够了。例如，为了测算行为人 A 和行为 B 之间的友谊，由任何一方知情者所提供的信息即可判断这种朋友关系的存在与否。也就是说，除非 A 和 B 的报告都缺失，否则我们就可以凭借一方的单独报告来判定友谊的存续。一般来说，对于一个含有 N 个行为人的非定向关系网络，当其中有 M 个行为人没有关系客体报告时，某一具体关系的应答率的计算如下所示：

$$R\begin{cases} = 100\% & \text{当} M = 0 \text{或} M = 1 \\ = \left(1 - \dfrac{C_M^2}{C_N^2}\right) \times 100\% & \text{当} 1 < M < N \\ = 0 & \text{当} M = N \end{cases}$$

例如，在一个含有五个行为人的网络中，节点应答率和关系应答率依关系节点缺失数量的不同而有所变化，具体见下表：

缺失节点数	节点应答率	关系应答率
0	100%	100%
1	80%	100%
2	60%	90%
3	40%	70%
4	20%	40%
5	0	0

为了方便阐述，现假设网络行为人分别为 A、B、C、D、E。在这五个行为人中的 10 种非定向对偶关系是 AB、AC、AD、AE、BC、BD、BE、CD、CE、DE。假如行为人 A 没有报告他/她的关系，关于 A 的这些对偶节点依然能够从其他四个人的报告中获得。因此，尽管缺失了 1 个节点的报告，

关系应答率仍然是100%。当节点缺失数量为2至4($1<M<N$)时,关系应答率是$\left(1-\frac{C_M^2}{C_N^2}\right)\times 100\%$。比如,假如有3个节点(A、B、C)未报告他们与其他人的关系时,关系应答率为70%$\left[\left(1-\frac{C_3^2}{C_5^2}\right)\times 100=70\%\right]$。在这种情况下,三对非定向性关系缺失了(AB、AC、BC),但是其他七对关系仍然可以由其他剩下的人报告。假如没有任何行为人提供信息,则节点应答率和关系应答率都降为0。因为节点应答率的计算公式为$\left(1-\frac{M}{N}\right)\times 100$,而且$\frac{M}{N}$总是要大于$\frac{C_M^2}{C_N^2}$(有关数学证明可以向本书作者索要),因此在节点缺失报告的每一个层级上,非定向关系网络的关系应答率总是会高于节点应答率。

相反,缺失的节点信息对于定向网络的关系应答率具有重大影响。不对称性存在于多种网络关系中,例如提供建议、信任、喜爱。行为人A和行为B可以构成两种定向性关系:AB,表示A所报告的其与B的关系;BA,表示B所报告的其与A的关系。因此,每一个节点的缺失,都会导致该节点所指向的所有关系信息的缺失。假定一个含有N个行为人的网络缺失了M个节点,那么该定向网络的关系应答率为:

$$R\begin{cases}=100\% & \text{当 } M=0\\ =\left(1-\frac{M\times(N-1)}{2\times C_N^2}\right)\times 100\% & \text{当 } 0<M<N\\ =0 & \text{当 } M=N\end{cases}$$

说明如下:现假定有一个五人的关系网络,它含有(5^2-

5)＝20种定向对偶关系：AB、BA、AC、CA、AD、DA、AE、EA、BC、CB、BD、DB、BE、EB、CD、DC、CE、EC、DE和ED。如果每个行为人都报告其与他人的关系，那我们就得到一个100%关系应答率。假如有一个点（比如A）缺失，那么所有A的报告也会缺失，导致这20对关系有4对（AB、AC、AD、AE）未报告，因此关系应答率就只有80%。在定向网络中的每一个点都必须评价与其他点的关系，因此M个缺失点（当$0<M<N$）将导致$M\times(N-1)$个关系的缺失。当然，当没有一个点报告关系时（$M=N$），节点应答率和关系应答率都等于0。

利用一点代数推理，我们现在证明在定向网络中关系应答率总是等于节点应答率。假设我们有一个含N个节点、M个缺失点的定向网络，那么节点应答率是$\left(1-\frac{M}{N}\right)$，而关系应答率为$\left(1-\frac{M\times(N-1)}{2\times C_N^2}\right)$。具体来说，$\frac{M\times(N-1)}{2\times C_N^2}=M\times(N-1)\div\left(2\times\frac{N!}{2!\times(N-2)!}\right)=\frac{M\times(N-1)}{N\times(N-1)}=\frac{M}{N}$。因此，在有5个节点的定向网络中，当没有节点缺失时，节点应答率和关系应答率都是100%；当1个节点缺失时，这两种应答率都为80%；当2个节点缺失时，为60%；当3个节点缺失时，为40%；当4个节点缺失时，为20%；当所有节点都缺失时，为0。

在前面的论述中，我们认为，在非定向网络中，对偶关系中的某个成员所进行的报告可以视为是对该对偶关系的可信测量。但学者一般都认为，这一观点必须谨慎采用，因为

经常发生单向关系(Stork & Richards, 1992)。例如,在一个交流网络中,A报告说与B进行交谈,但B报告说根本未与A交流。这种矛盾对下列做法提出了质疑,即仅从未缺席方提供的信息,来构建行为应答方与行为缺席方的关系。斯托克和理查兹(Stork & Richards, 1992)建议,关系双方特征的相似性,比如他们的年龄、性别和教育背景等,可作为认定他们对偶关系的一个很好的指标。另外,网络信度——由双方关系成员共同描述的所有对偶关系的比例——表明了一个单独的报告是在何种程度上揭示调查应答者与未应答者之间的联系。当网络信度高,并且关系双方具有相同的或相似的个人特征时,一方的报告可以作为他们关系的一个可信代表。科斯尼特(Kossinets, 2003)赞同斯托克和理查兹的方法,即当调查未应答者数量很少时,可以根据应答者的报告来构建对偶关系网络。

自我中心网络研究对数据缺失比较敏感,因为人们常常疏于描述关系客体间的联系。在研究1985年GSS讨论网络时,伯特(1987)的报告指出,GSS自我中心网络研究中的数据缺失并不严重:共有1534个调查回应者,他们列举了4483个讨论对象,其中只有66个应答者提供的网络数据不完整,涉及195个讨论对象。伯特比较了两种讨论对象——一种具有全部网络关系数据,另一种则只有部分网络关系数据。他发现前者关系客体的35%都与其他关系客体关系不紧密,这正如调查应答者所报告的那样,而后者的这个值则升至59%。这其中主要的含义是,在一个自我中心关系网络中,关系客体间所缺失的关系一般都是弱关系。关系主体似乎更倾向于报告那些与其他关系客体存在联系的客体关系,而

不是那些看起来比较孤立的客体关系。这一发现也解释了与其他研究方法结论相反的一个现象:受过大学教育的GSS调查回应者,相比教育程度较低的回应者,往往会提供更加不全面的自我中心网络关系数据。大学毕业生倾向于列举那些由许多弱联系的关系客体所构成的自我中心网络,反之也将不会更多地报告有关客体关系的数据缺失。

由于没有万无一失的事后补救办法来解决数据缺失问题,因此这个问题的解决之道只能是让更多的调查对象了解参与的重要性(Knoke & Kuklinski, 1982:35)。为了获得较高的参与程度,比如有些网络研究达到了90%的应答率,研究人员常常要付出更多的辛苦,采取不同的说服技巧,比如个人书信、电话联络以及钱财激励等。数据缺失问题这一难题并非仅仅存在于使用调查方法进行的社会网络研究中,档案研究也容易受数据缺失的影响。一般人会想当然地认为,档案数据的完整性是受数据存储管理操作的影响的。然而,在某类数据的可得性不变的情况下,编码人员的能力与数据搜索软件的效率也同样能对获取信息的数量和质量产生重大影响。训练较差的编码员、设计糟糕的软件,都绝对会导致规模较大的数据缺失。



第4章

网络分析基本方法

这一章我们将讨论社会网络分析的基本方法，对重要的传统论题和新近的发展也会给予同等的关注。其中，我们尤其着重论述密度、中心度、内聚、结构等价、聚类、多维度标度、块模型。由于网络数据描述总是在社会网络分析之前，因此这部分自然也就从描述两种独立但联系紧密的展示和分析社会网络数据的方法开始：图表和矩阵。图表展示的是社会网络的可视性，而矩阵则利用数学上的代数来描述网络关系。虽然社会网络学者可以自由地选择运用图表或矩阵模式来描述他们获得的数据，但是这两种方法各有优劣。图表可以为读者提供有关网络结构的更加直观的表述，但缺乏数学上的可操作性。与之对应的是，虽然矩阵看上去界面不够亲近读者，但他们非常便于对社会网络数据进行复杂的数学和计算机分析。网络分析者必须熟练运用这两种日常分析工具，以便对关系数据的重要特征进行概念化和测量。

第1节 | 图表

雅克布·莫雷诺(1934)是社会计量方法的先驱,他重视建立社会关系图,即二维图,用于展现某一有界社会系统内行为人之间的关系,比如一个小学班级。点和线——社会关系图中的两个基本元素——也确定了图的基本构成,从而也从图论中引入了相关概念和定理(Harary, 1969)。在社会关系图和图表中,行为人是由 N 个点(也称为节点或顶点)来代表的,常常用标识名、字母或数字来表示。两点之间划出的线(也称为弧或边),表示一种关系或连接;线的缺失,意味着两个行为人之间不存在直接关系。如果两点之间有一条线,那么这两点就是邻接的。如果某一个点是定义某条线的两点之中的一点,那么这个点就附属于这条线;反之亦然。在网络关系量表中,点的精确位置和线的长度是可以随意的,虽然某些版本要比另外的版本更清晰。建构一个富有洞察力的社会关系图,和进行一项科学活动一样,是一种艺术行为。

假如一条线没有箭头,那么就表示这种关系是非定向性的或双向的(如同事关系)。假如关系是从一个行为人指向另一行为人,那结果就是一个定向图表或有向图。形式上,一个有向图是点的有限集合,也是有序配对(a, b)的集合,其

中点 a 是线的起点,点 b 是线的终点。单向箭头表示一种定向关系,方向由箭尾的行为人指向箭头的行为人(如,提供建议)。双向箭头线条表示两个定向性联系,从每一个点到另一个点,表达的是一种互惠关系或相互关系(如,每个行为人都选择对方作为自己的"亲密朋友")。一个替代模式是使用两个单向箭头,每个箭头都指向对方。一些社会关系图也通过线条粗细的变化,或者不同种类破折号的使用,以便更形象地表达双方关系的力度、密度或者频率。

为了阐述一个社会网络结构的图表,我们利用了费尔德曼—萨维斯伯格等人(2005)对喀麦隆妇女同乡会组织的网络研究。为了便于分析集体记忆是如何影响妇女们对生育的讨论,费尔德曼—萨维斯伯格等人(2005)访谈了喀麦隆首都雅温德的六个妇女协会中的 156 个成员。他们的深度访谈涉及妇女社会网络的一些问题,比如"请根据下列标准对你和你所在协会的其他妇女的联系力度进行排序:(1)知己;(2)朋友;(3)认识;(4)完全陌生"。由于每一个妇女仅被问及她与同一协会其他妇女的关系排序,因此对这六个协会都分别构建了独立的网络结构。为了简单阐述,我们选取了第六协会的网络结构,它有六个协会成员。

图 4.1 表示第六协会成员间的社会关系图。线条的粗细表示三种层级的选择,没有连线表示两人之间完全陌生。在这六个成员间的 23 种定向关系中,只有 15 对关系中的 4 对是互相同意的(即两个妇女都选择了同样的层级):妇女 1 和妇女 2 互相视为知己,妇女 3 和妇女 5 是朋友,妇女 2 和妇女 5 仅仅认识,而妇女 5 和妇女 6 则完全陌生。其他 11 对则对她们的关系有不同的分级。例如,妇女 6 将妇女 2 视为知

己，但妇女2认为仅认识妇女6而已。虽然社会关系图可以提供一个直观的网络结构图，但是在展示由几十个甚至上百个行为人组成的大型网络方面，它们在视觉效果上就常显得很难看了。我们可以看出，仅仅六个行为人和23种定向关系，图表就显得非常复杂了。相比较而言，一个矩阵就可以轻而易举地表示几十个行为人的网络关系。

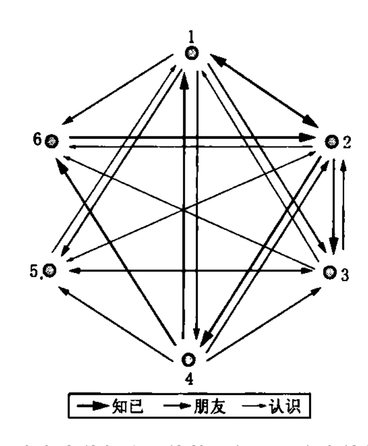


图4.1 喀麦隆首都雅温德第六妇女同乡会的社会网络

图4.2是一个包含七个虚拟行为人的有向图，表示的是提供建议的社会关系。通过这张图可以阐述几个基本的概念。子图，是指在一张大图中点和线的子集。汤姆(Tom)，贝蒂(Betty)，简(Jane)三人组成的三角关系便是一个子图。行程是附属的点和线交替的过程，在这个过程中，每个节点附属于它前后的线。行程的长度是它所包含的线的数量。行程简—汤姆—贝蒂—哈里(Harry)—迪克(Dick)—萨莉(Sally)的长度是5。轨迹是由不同的线构成的行程，即使其中一些点也许被访问过不止一次，例如迪克—萨莉—哈里—汤姆—贝蒂—哈里—迪克便是一个长度为6的轨迹。

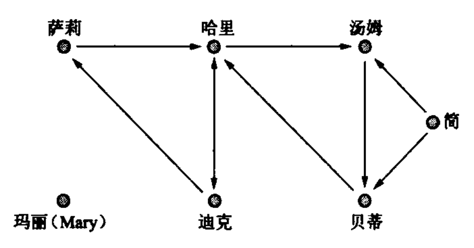


图 4.2 由七个行为人组成的有向社会网络

由于这个轨迹的起始和终止都在同一个点,因此它也被称为回路。沃瑟曼和浮士德(1994:107—108)讨论了封闭行程、旅程和回路等相关概念。

在网络分析中,一个重要概念是路径。它被定义为一个经由完全不同的节点和线的行程,其中点和线只能经过一次。在有向图中,一条路径必须经由不同的节点,而且其中的每条线都是由第一个节点到第二个节点。换句话说,所有的箭头都是朝着一个方向的。因此,贝蒂—哈里—萨莉不是一条路径,但贝蒂—哈里—迪克—萨莉是一条路径(见图 4.2)。路径的长度(即其行程中所经线的条数)被称为径长。在两个点之间也许有好几条路径,但是最短的路径被称为测地线。从简到萨莉的两条路径中,测地线简—贝蒂—哈里—迪克—萨莉的径长为 4,它短于另一条长度为 5 的路径(简—汤姆—贝蒂—哈里—迪克—萨莉)。那从迪克到汤姆的测地线是多少呢?

如果两个节点之间存在至少一条任何长度的路径,那么这两点就是可及的。汤姆对萨莉来说是可及的;虽然对萨莉来说,简是不可及的,但是对简来说萨莉是可及的。

如果每对节点之间存在路径,那么这个图表就是连接状态的;但如果至少有一对节点之间没有路径,那这

个图标就是非连接状态的。如果一个点没有与任何其他点有线连接，那这个点就是孤立的。玛丽明显是孤立于其他六个人的，因此这个图也是处于非连接状态的。假如我们忽略玛丽，那剩余的六人关系图就展示了三个层次的连接：

强连接：每两个点都由直接的路径双向连接（例如，从 a 到 b 并且还从 b 到 a）；

单向连接：每两点之间都从一个方向（而非另外方向）由一条路径相连接；

弱连接：所有两点之间有连线，但不考虑其连接方向。

很明显，这个六人的有向图只是个弱连接图。然而，在大网络之中的几个小网络子集的连接强度要比整个网络图强。图表组成部分是构成一个连接状态图表的极大化子图，而一个强连接图组成部分就是构成强连接图表的极大化子图。子图萨莉—哈里—汤姆—贝蒂是一个单项连接部分，因为不同长度的定向性路径单方向地连接了每一对点。这个组成部分是极大化的，因为一旦加入简，就将破坏这种单向联系。那为什么简—汤姆—贝蒂这一更小的子图不属于强连接部分或单向连接部分呢?

在一个连接图中，如果去除一个点，将导致该图不能连接（即会创建两个或多个组成部分），那这个点就是一个分界点。假如将哈里从图 4.2 中移除，那这个六点图就会分成两个小的组成部分（萨莉—迪克和汤姆—贝蒂—简）。同样的，如果去除一条线，会将一个图表分成两个或多个组成部分，那这条线就称为桥。请问这张图中包含桥么？为什么?

相对于存在很多维持信息和资源流的多余路径的网络，有分界点和桥的网络更容易受到破坏。例如，恐怖基地的秘密网络有可能因为一个关键人物被抓或被杀而陷入瘫痪状态(Borgatti, 2003; Tsvetovat & Carley, 2005)。

第 2 节 | 矩阵

网络关系的代数形式可以表达社会关系图中所有的量化信息，和直观的图表形式相比，代数形式能够做更多的分析。可供数学分析的社会网络数据的最基本形式是列表，称为社会矩阵，也称为邻接矩阵，是一个由行和列所构成的数值方列阵。例如，大写黑体的表达式 $\mathbf{X}(N, N)$表示社会矩阵 $\mathbf{X}$，它有 N 个行和列，而且其 N^2 个单元表示 N 个社会行为人之间的各种关系。行为人的顺序或定位可以通过矩阵的行和列来一一对应，利用下标字母 i 和 j（其值从 1 至 N），相应表示特定的行列位置。大体上，一个单位格的值表示为 x_{ij}；例如，x_{35}表示一个社会矩阵中第 3 行第 5 列的值。在绝大多数社会网络的实际运用中，矩阵主对角线（即第 i 行第 i 列所组成的单位格）是没有意义的，例如，某个人不可能是他/她自己的好朋友，因此在数据分析中是可以被忽略的。因此，在一个含有 N 个行为人的社会矩阵里，拥有唯一值的单元格的最大值是 $N^2 - N$。

在一个社会矩阵单元格的数值测量矩阵中，特定的行和列对应一对行为人之间的特定关系。按照惯例，在定向关系中，行中的行为人属于起点或发送者，而列中的行为人属于终点或关系的接受者。最基本的测量就是关系的存在或缺

失,分别由 1 和 0 这两个二进制值表示。因此,二进制矩阵数 $x_{ij}=1$ 表示行为人 i 发送了一个关系给行为人 j;如果 $x_{ij}=0$,则表示由 i 至 j 没有关系。由于在有向图中,关系并不局限于双向的或者互惠的性质,因此它们的矩阵图是不对称的。在非定向网络中,$x_{ij}=x_{ji}$,也就是说,起点 i 和终点 j 之间的关系值总是等同于起点 j 和终点 i 之间的关系值。这种矩阵是对称的。

社会矩阵也包含非二进制值,反映关系强度,比如联系的频率、关系的力度或者组织的大小。在这类赋值的图表中,单元格可以从 0 至对偶关系水平的极大值,或者包括小数甚至负数以表示反向关系,比如“不喜欢”或者“敌对”。

表 4.1 给出了喀麦隆雅温德地区第六妇女同乡会的矩阵图。这个矩阵是图 4.1 的代数形式。由于每一位妇女都对与其他五位妇女的关系力度进行打分,并且在双方评价时会有差异,因此这个矩阵是赋值和不对称的。行中的妇女对与其他妇女的关系进行评价,而列中的妇女是被打分的对象(行和列中的标注并不是矩阵的一部分,但是可以作为参考)。例如,妇女 1 对与妇女 3 的关系力度评分为 2(表示朋

表 4.1 喀麦隆雅温德地区第六妇女同乡会的社会网络矩阵

	W1	W2	W3	W4	W5	W6
W1	—	3	2	2	2	2
W2	3	—	3	3	1	1
W3	1	2	—	0	2	1
W4	3	2	2	—	2	3
W5	1	1	2	0	—	0
W6	0	3	0	0	0	—

注:0 = 完全陌生;1 = 认识;2 = 朋友;3 = 知己。

友关系),但是妇女3对她与妇女1的关系的评分只有1(表示认识)。图中的破折号表示主对角线上的值是没有意义的。表4.2是图4.2的一个二进制赋值矩阵。

表4.2 7人有向社会网络(图4.2)的矩阵图

	萨莉	哈里	迪克	汤姆	贝蒂	简	玛丽
萨莉	—	1	0	0	0	0	0
哈里	0	—	1	1	0	0	0
迪克	1	1	—	0	0	0	0
汤姆	0	0	0	—	1	0	0
贝蒂	0	1	0	0	—	0	0
简	0	0	0	1	1	—	0
玛丽	0	0	0	0	0	0	—

有时,网络研究人员使用非正方形矩阵表示行为人特征或者他们对某一事件的参与。数学符号 $Z(N, M)$表示一个长方形矩阵 Z,其中 N 是行为人的数量,M 是特征、事件或地址的数量。例如,弗里曼和韦伯斯特(1994)观察了31天内43个沙滩常客,记录了与这些人有关的353个事件(比如野餐、游戏等)。他们建立了一个43×353的矩阵,其中第 i 行第 j 列的值为1,表示第 i 个人参与了事件 j。如果研究人员收集了不止一种关系的网络数据,那么可以分开或综合构建并分析多个社会矩阵。表示 R 多元网络间具体关系的矩阵形式,只需要简单地多加个脚注,也就是 x_{ijk},表示在第 k 种关系中第 i 行和第 j 列的单元值。

第3节 | 关系的测量方法

社会网络数据中的基本关系测量方法包括密度和中心度。在讨论它们之前,我们回顾几个网络数据的基本原则,它们在向度和赋值这两个维度上有所区别。向度将社会网络数据分为无向关系和有向关系两种。在无向关系数据中,一对关系的起点和终点(例如,发送者和接受者,选择者和被选者)是没有区别的:如果迪克娶了简,那简也同样嫁给了迪克。其他种类的无向连接包括某一组织内成员关系,某一社会事件的共同参与,以及互相交流知识等。为了获取无向关系资料,研究者常常只需要问一方知情者,然后利用这个应答即可了解这个对偶连接。相反,当对偶关系的每一方都向对方启动一种关系时,有向关系发生了,结果导致四种可能的二进制组合:(1)A给B一些东西,比如信息、建议和钱等,但B不给A;(2)B给A,但A不给B;(3)A和B之间互相交换;(4)A和B之间没有交易。有向数据经常收集到这样的情形:关系主体的自我报告认为存在某种网络关系,但这种关系与关系客体对其看法不一致。许多也许在理论上可以被归为无向关系的网络数据,在实际中应当视为是有向连接数据来测量。例如,为评估一项信任关系的存在,研究人员应当问关系双方是否信任对方。研究人员不应首先假定

朋友关系是否总是互惠的，而应首先将他们视为是内在的有向关系：Ted也许视Sally为朋友，但Sally可不认为Ted是她的朋友。正如第3章所讲到的，许多的认知社会结构网络关系具有固有的有向性，因为它们是来自于行为人对主观信仰、态度或者关系认识的报告，而这些很可能是不同于关系客体的主观看法的。即使一个关系在理论上可以归为无向类，但研究者通常还是得选择谨慎一点的做法，即请关系双方都提出各自的报告。在随后的分析中，研究者可以利用弱标准（比如只要一方报告认同，即认为关系存在），或强标准（要求双方都必须确认该关系）来整合这种不同。一旦采用这一观点，那么在数据收集上就能使得分析者在供替代的标准假设之下验证他们结果的稳健性。

根据关系信息的可获得性，社会网络联系可以分为二进制的或赋值的量表。在二进制数据中，被试者只需要回应关系的存在（编码为1）或缺失（编码为0）。在赋值数据中，他们不但要表示关系的存在与否，还要对二者交往的强度或频率进行排序或量化。例如，青少年会被问及在去年与伙伴发生性关系的次数。因此，二进制网络数据只有两个值，0或1。但赋值关系数据则展示了一个有序的或连续的梯度，数值可以从0至极大的可能值，甚至可以涵盖小数或分数。

将向度和值进行交叉列表，可以得到四种基本的社会网络数据类型：二进制无向、二进制有向、赋值无向、赋值有向。二进制无向数据需要从受访者或知情人那里获得的信息最少。访谈者只需要询问一方即可确认关系的存在与否。赋值无向关系数据需要一方来量化双方的关系。二进制有向测量只需要双方提及关系的存在与否。但是赋值有向数据

需要最详尽的信息,即关于对偶关系重要性的两个报告。由于自我中心网络研究和全幅网络研究在它们的研究设计、抽样、数据分析及其解释等方面存在极大的差异,因此下面两个小节讨论适用于这四类网络数据的测量。

自我中心网络的测量

在自我中心网络研究中,知情人被要求报告与其有特别联系的关系客体。这种关系被明确表述在提供给被调查者的提名法则中。一个著名的例子就是美国综合社会调查,"讨论重要问题"的提名法,在第三部分有论述。没有独立的信息是从由 1985 年 GSS 被调查者所提名的任何关系客体那里收集的。这一小节主要集中介绍自我中心网络的密度和网络范围测量。

密度是指测量自我中心网络中 N 个客体之间互相联络的程度(忽略了关系主体 ego,因为本来所有的客体都与之直接联系)。现假定这种关系是无向性的二进制(有/无)关系,那关系密度(D)就等于所报告的客体间对偶联系(用 L 表示)数除以此种关系的极大可能数:

$$D = \frac{L}{C_N^2}$$

其中 $C_N^2 = \frac{N!}{2! \times (N-2)!}$。例如,假设主体有五个关系客体,并且客体间只报告了两种关系,由于 $C_5^2 = \frac{5!}{2! \times (5-2)!} = \frac{(1 \times 2 \times 3 \times 4 \times 5)}{(1 \times 2) \times (1 \times 2 \times 3)} = \frac{120}{12} = 10$,那么这个社会网络的关

系密度就是 $D=\frac{L}{C_N^2}=\frac{2}{10}=0.20$ 或 20%。

如果一个自我中心网络的联系是根据有向二进制来测量的，那么密度计算公式就为：

$$D=\frac{L}{2\times C_N^2}$$

上述两种密度计算公式的第一个区别是分母不同。在同一 N 值的社会网络中，有向连接数是无向连接数的两倍。第二个区别是分子不同。虽然两个公式的分子都是反映对偶联系的真实呈现数量，但后者必须检查两个关系方向。在二进制无向网络数据中，每对连接方只计算一次，但在二进制有向网络数据中，每对连接方被计算两次，因为关系主体要报告每个客体是否首先与所有其他客体方进行联系。

接下来，假定网络数据是赋值的无向性关系，那么自我中心网络密度计算如下：

$$D=\frac{\sum L_w}{C_N^2}$$

在这个公式中，分子表示各种关系的所有值的和。此分子除以 N 个行为人间可能发生的所有关系数，便得到关系密度值，反映了主体之客体间关系的平均力度。假设某一研究项目要求主体对其 5 个关系客体(A、B、C、D 和 E)进行排序，标准如下：无关系(0)、认识(1)、朋友(2)、密友(3)。假如主体将 A 视为 B 的密友，将 D 视为 E 的朋友，而其余人之间没有关系，那么密度 $=(2+3)/10=0.5$，结果表明这个自我中心网络的 10 个关系呈现的平均力度或权重为 0.5。

最后一种密度测量是关于赋值有向关系数据的,公式如下:

$$D=\frac{\sum L_w}{2\times C_N^2}$$

它的分子包括所呈现关系的所有赋值的总和,而分母则包括 N 个行为人所在网络的所有可能有向关系数量的最大值,这是同样规模网络无向关系数的两倍。

二进制网络数据的密度测量值被限制在 0 和 1 之间,其中 0 表示关系客体间没有任何联系,1 表示所有的客体间都存在直接互相联系。这种限制的存在是因为二进制网络密度的分子不可能大于分母。对于无向和有向二进制数据二者来说,双方实际关系的数量必然要小于或等于所有可能的对偶关系的最大数(当 L 仅表示关系的存在数量时,$L\leqslant C_N^2$;当 L 不仅考虑关系的存在,还考虑关系的方向时,$L\leqslant 2\times C_N^2$)。然而,对赋值的自我中心网络数据来说,密度能够超过 1.00,因为对关系特征的赋值也许远远大于二进制数据的最大关系值 1。

网络范围在广义上被定义为一个行为人的网络与其他不同行为人的联系的程度。在实践中,研究者利用多种经验标准来测量网络范围,例如网络规模、密度和多样性(Campbell, Marsden & Hurlbert, 1986)。而且,自我中心网络多样性测量的是一个个体网络中客体社会特征异质的程度。依据客体特征的测量水平,自我中心网络密度可以通过连续变量的标准差来进行测量,或者通过分类变量(包括离散的或有序的)的定性变化指数来测量。

当一个客体特征按照连续变量来测量时,标准差反映的

就是一个自我中心网络内特征的差异性。一般来说，对含有N个客体的第i个主体，其中第j个客体的特征X_{ij}是一个连续变量，那么标准差就是：

$$s_{X_i}=\sqrt{\frac{\sum_{J=1}^{N}(X_{ij}-\bar{X})^2}{N-1}}$$

所有主体标准差的平均值可以测量一个自我中心网络样本中特征的多样性。马斯登(1987)发现，1985年GSS调查所得客体教育年限的平均标准差只有1.37年，这意味着在美国人的核心讨论圈中讨论对象都倾向于拥有相似的教育水平。

当客体特征是按照分类变量来测量时，那IQV就是用来测量自我中心网络的多样性。通常，对含有N个客体的第i个主体而言，其中所有客体都被分成K类离散的或有序的类别，那IQV等于：

$$IQV_i=\frac{1-\sum_{J=1}^{K}p_j^2}{(K-1)/K}$$

其中p_j^2是第j类中客体的比例。IQV的标准值介于0.00和1.00之间，其中0.00表示所有的N种情形都包含在一个类别之中，而1.00表示所有的客体都均匀分布在K个类别内。例如，假如某主体提名了4个男性和1个女性作为关系客体，那性别构成IQV就是$\frac{1-[(0.8)^2+(0.2)^2]}{(2-1)/2}=0.64$；当某人提名3个男性和2个女性时，那得到的性别构成IQV就是0.96。因此，后一自我中心网络展现了更大的性别多样性。所有关系主体的IQV均值构成了一个样本定性多样性的分类变量指数。利用这一方法，马斯登(1987)报告说美国

人自我中心网络的性别 IQV 是 0.68,而种族 IQV 只有0.05。很明显,美国人的核心讨论圈在性别构成上要比在种族差异上均衡得多。

全幅网络测量

在自我中心网络研究中,一个中心个体通过列举与他/她有特定关系的客体来构建其网络。除非从一个高密度社区抽样——其中关系主体和客体之间都联系紧密(White & Watkins, 2000),否则某一个主体的关系客体极少与另一主体的关系客体有关系。因此,之前部分中讨论的密度测量是建立在互无联系的知情人所提出的关系客体的特征基础之上的。相反,在全幅网络研究中,研究人员设定一个网络边界,只关注在此范围内的行为人之间的联系。例如,克雷克哈德(1987)的研究关注某一组织内 21 位高级技术经理人。每一位都被要求确认是否从其他 20 位经理人那里获得过建议。利用全幅网络设计获得的数据,只由所选出的这些行为人的联结构成。通常,选定范围之外的行为人的联结是不会考虑的。因此,把任何一种自我中心网络密度测量法应用到全幅网络中只得到单个的值。N 个行为人的随机抽样,产生 N 个自我中心网络密度值,因为每一个行为人的自我中心网络都不同于其他行为人。相反,全幅网络有一个群体的单个密度值,不论总体 N 是多少。尽管在分析层次上有这种差异,但对于自我中心网络和全幅网络而言,这四种密度测量的计算公式都是相同的。

除密度之外,可广泛适用于全幅网络的其他关系测量方

法还包括节点度、行程、轨迹、路径、回路、组成部分、测地距、直径、分界点、桥、树、同构子图等。本书无法做到对所有这些方法进行详尽阐述，相关资料可见沃瑟曼和浮士德(1994)。这里，我们讨论两个基本主题：节点度和测地线。这两种测量方法同样都可以用于全幅网络和自我中心网络。为方便讲解，我们以代表不同的网络数据类型的四个图标为例。

节点度表示为 $d(N_i)$，是行为人 i（也就是节点）的所有关系数量，其中度指线的条数。在一个二进制无向图中，如图 4.3所示，每一个行为人的节点度为 $d(N_A) = 4$，$d(N_B) = d(N_C) = d(N_D) = d(N_E) = 3$。该全幅网络的平均节点度，由所有行为人的节点度相加，再除以行为人总数，表示节点间的联结程度。对图 4.3 来说，平均节点度是 $\frac{4+3+3+3+3}{5} = 3.2$。

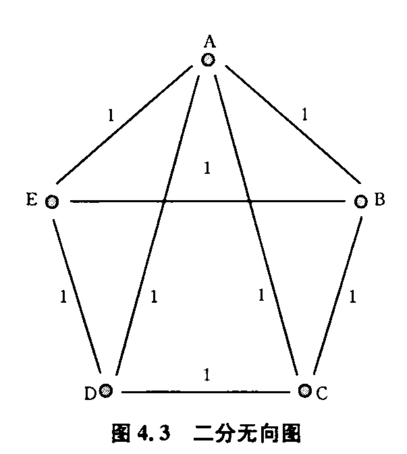


图 4.3　二分无向图

然而对于有向图，研究者必须区分节点入度和节点出度。节点入度是一个行为人接受的从其他行为人方向所来的线的数量；节点出度是从一个行为人节点向其他行为人所发出的线的数量。图 4.4是一个朋友提名网络，其中 A 提名

B和E作为朋友,但并没有其他人将A作为朋友。因此,A的节点入度为0,但节点出度为2。虽然一个人的节点入度和出度会有差别,但全幅网络的平均节点入度和出度总是相等的,因为指向所有行为人的线条总数等于从所有行为人处发出的线条总数。在图4.4中,行为人A、B、C、D、E的节点入度分别为0、2、1、1、3,但它们的节点出度为2、1、1、2、1。每一套度数总和都是7,因此网络平均节点度数为$\frac{7}{5}=1.4$。一些学者认为,一般意义上节点出度表示扩展程度,而节点入度代表接纳程度或者受欢迎程度(Wasserman & Faust, 1994:126)。根据这种有趣的观点,A是一个开朗型的行为人——会选择他人作为朋友,但是他是一个在网络中最不受欢迎的人。相反,E是最受欢迎的,但相对来说他是个内向型的人。

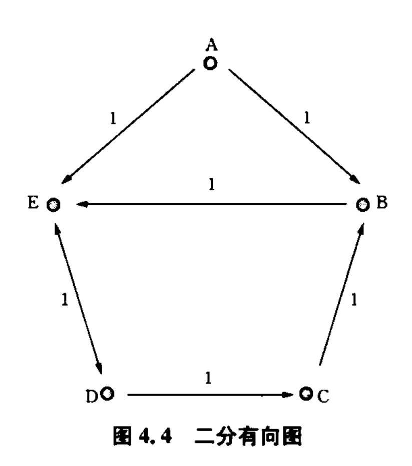


图4.4 二分有向图

对于计算赋值图的节点度数,通过网络文献回顾可以得到些许启示。其中一个简化的解决办法,就是将赋值图中所有的非零值替换为1,然后采用二进制图的处理办法。但是,

这一方法忽略了赋值图中有关联结力度的丰富信息。在此，我们推荐一个概括的办法，用于计算赋值的无向图的节点度数。这一方法将一个行为人所有线条的赋值加起来。然后，该赋值的无向图的平均节点度数就是所有赋值节点度数的总和除以行为人的总数。图 4.5 中 A、B、C、D、E 的节点度数分别为 5、13、7、5、12。因此该图的平均节点度数是 $\frac{5+13+7+5+12}{5}=8.4$。我们承认，这一求和法在下述两种相反的情形之下，会产生较高的节点度数：行为人线条值低，但数量多；行为人线条数量少，但赋值高。例如，有一朋友关系网络，其中的线条赋值表示朋友关系节点的力度。第一种情形下，一个人有许多关系浅的朋友；第二种情形下，一个人只有少许关系非常密切的朋友。概括方法为两种节点度数都提供了同样的高值，但是并不能将二者区分开来。

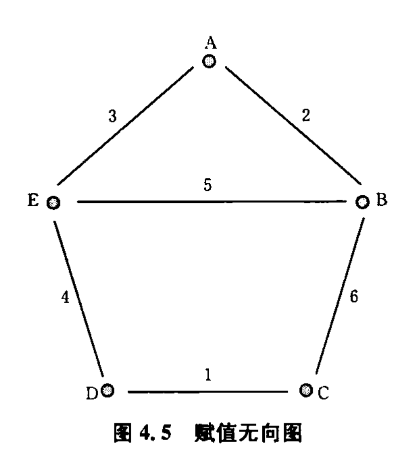


图 4.5 赋值无向图

解决这个难题的一个办法，我们称之为均值法，它将赋值的线条数量考虑进来，用某一行为人线条的赋值总数除以线条总数。因此，均值法计算与行为人直接联系的线条的平均赋值。

这一解决办法可以区分前述段落中所论述的两种情形。

假定行为人 X 有 5 个关系浅的朋友，每个朋友关系赋值为 1；行为人 Y 只有 2 个朋友，但每个朋友关系都高赋值，分别为 2 和 3。求和法得到同样的节点度数 5。而均值法则各产生一个平均数，行为人 X 的节点平均数为 $\frac{1+1+1+1+1}{5}=1$，而 Y 的平均数为 $\frac{2+3}{2}=2.5$。Y 的平均节点度高于 X 表明 Y 比 X 有着更为紧密的朋友关系。

对赋值有向图的节点度数测量在使用节点入度或节点出度时，必须考虑关系的方向性。下表给出了利用求和法和均值法获得有 5 个关系节点的图 4.6 中节点入度和节点出度的结果。

	节点度	A	B	C	D	E
求和法	节点入度	0	8	1	4	12
	节点出度	5	5	6	5	4
均值法	节点入度	0	4	1	4	4
	节点出度	2.5	5	6	2.5	4

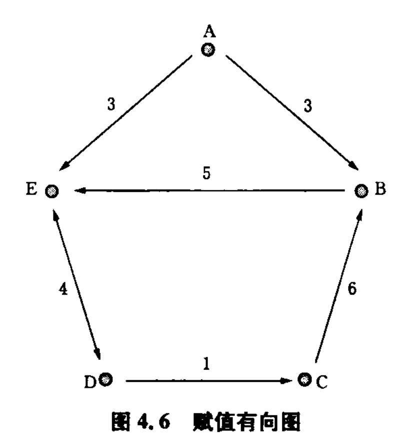


图 4.6 赋值有向图

对两种方法进行比较，可以发现几点有趣的现象。使用均值法，节点B、D、E有相同的节点入度，表明这三者同样受欢迎。相反，求和法则发现三个节点入度成较陡的层次区别，其中E最受欢迎，B次之，D位于最末。如图4.6所示，求和法显示E得到程度不等的最多提名，因此将E视为是最受欢迎的点。均值法强调，虽然D只得到一次提名，但是它有较高的赋值，因此D的排名与E和B一样。根据他们不同的研究目的，网络研究者可以选择不同的测量方法来测量有向赋值图的节点度，但必须谨慎解释他们的结果，解释的时候必须和求这些值的方式一致。

测地距是两个行为人之间最短路径的长度。它的计算因全幅网络的四种类型不同而差别较大。让我们从最简单的例子开始：图4.3的二进制无向图。节点A和B有许多条路径相连（例如，AB，ACB，AEB，ADCB，ACDEB，AEDCB），路径长度从1至4。最短路径（AB）的长度为1，因此A和B之间的测地距是1。图4.3包括10对关系（$C_N^2=\frac{5!}{2!\times 3!}=10$）。这10对关系间的测地距，依两点之间是否有直接连接而距离不同。例如，AB的测地距是1，而CE的测地距是2（通过CDE，CBE或CAE）。

含有N个行为人的有向图的双向连接数是同等规模的无向图的两倍，因为在有向图中，每对行为人有两个有方向的连接，而在无向图中则只有一对连接。例如，图4.4中AB之间有两个有向关系：AB和BA。由A连至B，有两个路径：AB和AEDCB（但AEB并不是一条路径）。最短路径是AB，因此AB间的测地距是1。反过来，B不能连接至A，因此BA

之间的测地距不明确。有向图的测地距分析,在网络研究中有许多实际的运用。其中一种情形,就是绘制一个行为人群体之间的交流图。假定图 4.4 表现了一个五人间的交流网络,A 可以直接向 B 发送信息,或者通过 E、D、C 向 B 间接地发送信息。但是通过中介传送信息容易延误或者歪曲,因此正如 AB 间的测地距所分析的那样,A 最有可能直接与 B 交流。相反,B 既不能直接也不能间接地向 A 发送信息。实际上,A 只能充当一个信息发送者,而不能作为一个信息接收者,因为没有人与 A 交流。这个例子生动地描述了信息的流动并非是互惠的。

当考虑每一条路径的赋值时,计算赋值有向图的测地距会面临一些困难。我们可以将赋值图简化成二进制图,然后运用二进制图处理程序,但因此也浪费了这些赋值所提供的丰富信息。杨和诺克(2001)提出了一个解决之道:利用路径长度和路径值来测量赋值图双方的测地距。这一方法通过一个合适的运算法则来实行(Yang & Hexmoor, 2004)。大体上,这个方法涵盖:(1)找出所有的连接路径;(2)找出每一条路径的最小值;(3)用最小值除以每一条路径的长度;(4)在第三步的结果中挑选出最高的值,这便是一个赋值图的测地距(也称为最佳连接,Yang & Knoke, 2001)。这个方法假定,长的路径最容易受到延迟和扭曲,而最短的路径对于双向交流来说等同于瓶颈路段。在图 4.5 中,连接 A 和 B 有三条路径,他们的平均路径值为:$AB = 2$, $AEB = \frac{3}{2} = 1.5$, $AEDCB = \frac{1}{4} = 0.25$。选择最高的平均路径值,那 A 和 B 之间的测地距为 2,因此最佳路径就是 AB。

杨和诺克(2001)的方法可以运用于赋值有向图中两点间的测地距的计算，不过关系的方向必须被考虑其中。图4.6连接A到B的两条路径是AB和AEDCB，其中AB平均路径值是2，AEDCB的平均路径值是0.25。选择最大的路径值，那AB间的测地距就是2。然而，连接BA间的测地距是无法确定的，因为B无法通过任何路径到达A。

节点度和测地距是社会网络分析的重要测量方法。节点度比较容易计算，并且常能为实际运用提供大量信息。节点度一般表示一个行为人参与网络行为的程度。例如，朋友关系网络中的节点度低，表示该行为人朋友较少。对全幅网络而言，平均节点度表示行为的综合水平。例如，一个社区住户间串门的平均节点度较低，意味着邻居间活动少。节点度也是节点度中心度测量的一个基础，这个在下文中将要讨论到。两节点间的测地距是节点之间关系亲密度非常重要的标准。由于每对行为人间都有几条路径相连，因此测地距就是每两节点间的最短路径，也是基于距离的综合法的基础，例如多维标度和多层聚类分析，这些在本节的后面将会进行讨论。

第4节 | 中心度和声望

图论在社会网络分析中的一个主要作用是，在个人分析层次和群体分析层次上，找出其中重要的或突出的行为人。中心度和声望的概念及测量通过综合所有点之间的结构关系，试图将图论关于全幅网络中行为人知名度的理念进行量化。对知名度的分析，将考量社会计量学上的选择与被选择，有时也包括间接关系的结构。个人行为人的知名度反映了他/她在其他行为人当中较为明显的可见性。群体层次上的中心度和声望指数，评估所有行为人的知名度在一个网络中的分散程度或不平衡程度。在本部分中，我们讨论两个有关知名度测量标准的大致分类：(1)中心度，无论是作为关系的发起者还是接受者，当一个行为人知名度较高时，他/她对多种关系的参与程度较高；(2)声望，表示知名行为人虽然只发起较少的关系，但是接受很多直接关系。知名度测量的选择，取决于概念化的运用和可得网络数据的类型。它们的计算要求运用前面几节讨论的矩阵数学概念。现在我们从最简单的例子开始，先讨论无向图，接着再讨论有向图。由于绝大多数知名度指标只要求二进制联系测量，因此我们的讨论就限于这类数据，但是我们鼓励网络研究学者致力于建立赋值图的知名度测量方法。

最广泛使用的中心度测量标准是程度、亲密性和中间性(Freeman，1977、1979)。这些测量在运用于无向关系和有向关系时有所区别，在运用于个体行为人和群体或全幅网络层次时也有所区别。下面的计算公式利用了沃瑟曼和浮士德(1994)的表示方法。

程度中心度

在无向二进制图中，行为人程度中心度测量社会网络中一个节点与所有其他节点相联系的程度。对于一个拥有 g 个行为人的无向图，行为人 i 的程度中心度是 i 与其他 $g-1$ 个其他行为人的直接联系总数，用矩阵表示如下：

$$C_D(N_i) = \sum_{J=1}^{g} x_{ij} (i \neq j) \qquad [4.1]$$

其中 $C_D(N_i)$表示节点 i 的程度中心度，$\sum_{J=1}^{g} x_{ij}$ 用于计算节点 i 与其他 $g-1$ 个 j 节点($i \neq j$，排除 i 与自身的联系；也就是说，主对角线的值可以忽略)之间的直接联系的数量。$C_D(N_i)$的计算就是简单地将行为人 i 在社会关系矩阵中对应的行或列所在的单元格值加总(因为无向关系构成一个对称性数据矩阵，因此行和列相同的单元格的值相同)。

如此测量的行为人程度中心度，不仅反映了每个节点与其他节点的关联性，而且也视网络规模(g)而定。也就是说，网络规模越大，程度中心度的最大可能值就越高。因此，一项具体的程度中心度值可表示一个小网络内行为人的关系

紧密，也可以表示在一个大网络内行为人只与不多的几个人联系。为了消除网络规模变化对程度中心度的影响，沃瑟曼和浮士德(1994:179)提出了一个标准化的测量公式：

$$C'_D(N_i) = \frac{C_D(N_i)}{g-1} \qquad [4.2]$$

这个标准化行为人程度中心度测量，使用行为人 i 的程度中心度值除以与其他 $g-1$ 个行为人最大可能连接数，得到与行为人 i 有直接联系的网络成员的比例。

这个比例从 0.0 到 1.0，0.0 表示与任何行为人都没有联系(例如一个孤点)，1.0表示与每一个人都有直接联系。标准化的行为人程度中心度测量行为人在诸多关系中的参与程度。得到高分的行为人是网络中最显眼的参与者。尤其是，标准化程度中心度值越接近 1.00，那么行为人在关系网络中的参与度越高。假如研究者更看重行为的纯粹数量而不是无向关系的具体来源和目标的话，那么他们很容易将中心度这一概念运用于信息网络的获取、控制和居间行为的分析(Knoke & Burt，1983)。

与行为人的程度中心度不同，组程度中心度测量一个社会网络中的行为人与其他行为人在个人程度中心度方面的不同程度。组程度中心度非常接近描述统计学的离差测量，例如标准差，表示围绕中心趋势值变化或扩散的量。弗里曼(1979)提出了组程度中心度的通用测量方法：

$$C_A = \frac{\sum_{i=1}^{g}[C_A(N^*) - C_A(N_i)]}{\max\sum_{i=1}^{g}[C_A(N^*) - C_A(N_i)]} \qquad [4.3]$$

其中，$C_A(N^*)$表示网络中所观察到的行为人程度中心度的最大值，而$C_A(N_i)$是$g-1$个其他行为人的程度中心度。因此，分子是行为人最大中心度与其他所有人的中心度之差的总和。分母是这些差值之和的最大理论可能值。

按照弗里曼的通用测量方法，沃瑟曼和浮士德(1994：180)在他们的组程度中心度指数中提出了有关分母的处理方法：

$$C_D = \frac{\sum_{i=1}^{g}[C_D(N^*) - C_D(N_i)]}{(g-1)(g-2)} \qquad [4.4]$$

分子表示最大中心度的行为人与其他节点之间的程度中心度所观测到的差值之和，分母测量差值的最大可能总和。这个值可能发生在星形图中，其中一个节点与其他所有节点相连，但其他所有节点都只与第一个节点相连。因此第一个节点具有最高的可能程度中心度为$(g-1)$，然而其他每个节点的程度中心度都等于1，两者之间中心度的差值就是$(g-1)-1=g-2$。因为这个差值在该图中发生$(g-1)$次，因此分母值为$(g-1)(g-2)$。组程度中心度指数值在0.0至1.0之间。当社会网络的程度中心度均匀分布时，即每一个点都具有相同的程度中心度时，$\sum_{i=1}^{g}[C_D(N^*) - C_D(N_i)]$等于0；因此，组程度中心度为0。在另一极端情形下，当一个节点具有最高可能的中心度$(g-1)$，而其他所有节点的程度中心度为1时，表示最大的不均匀分布。分子等于分母，组程度中心度指数为1.0。因此，组程度中心度越接近1.0，社会网络中节点的程度中心度就越不均匀或越呈等级分化状。

亲密中心度

行为人亲密中心度是用于反映在一个社会网络中某一节点与其他节点之间的接近程度(Sabidussi, 1966)。亲密和距离指一个行为人能多快与他人联系上,例如通过直接交流或者极少的中间环节。行为人的亲密中心度是它与其他点之间测地距的函数(如前所述,测地距是连接两节点间的最短路径长度)。行为人亲密中心度指数为行为人 i 与 $g-1$ 个其他行为人之间测地距总和的倒数:

$$C_C(N_i) = \frac{1}{\left[\sum_{j=1}^{g} d(N_i, N_j)\right]} (i \neq j) \qquad [4.5]$$

这个测量值不可能为 0.0,因为在数学中 0 是不可能做分母的。因此,孤点(从节点 i 至该节点的距离是不可测的)亲密中心度值是不可能测量的,该指数只能用于可形成连接图的节点。对于实证社会网络分析来说,这一限制要求每一个行为人在网络连接图中至少要与其他一个成员存在连接关系。极低的行为人亲密中心度值,或中心行为人与其他人之间的高距离值,一般是因为该行为人处于一个较大的网络之中,或该行为人处于一个小网络中,但与该网络中的其他行为人有相对较长的距离值。

当某一行为人与所有其他人关系亲密时,即在网络中与每个人都有直接联系时, $C_C = 1/(g-1)$,该值随网络规模大小发生变化。

为了控制网络规模便于不同网络间行为人的比较,比彻

姆(Beauchamp, 1965)建议,通过乘以 $(g-1)$ 来标准化行为人亲密中心度指数:

$$C'_C(N_i)=(g-1)(C_C(N_i)) \qquad [4.6]$$

为了比较和对比亲密中心度与标准化亲密中心度指数,我们参考两个例子。图 4.7 展示了一个 3 节点网络结构图,其中节点 A 直接与节点 B 相连,测地距为 1;间接与节点 C 连接,测地距为 2。因此,图 4.7 中节点 A 的亲密中心度是 $1/(1+2)=1/3$。

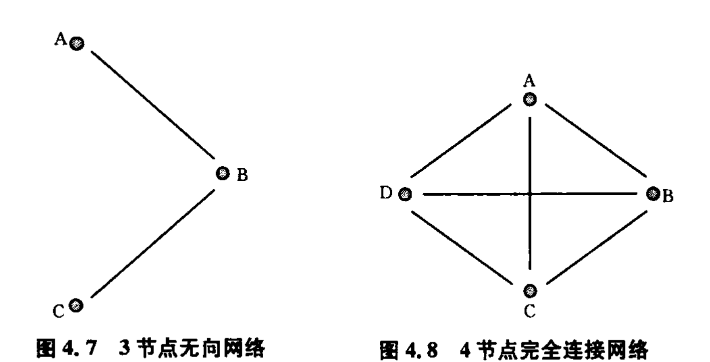


图 4.7 3 节点无向网络　　图 4.8 4 节点完全连接网络

图 4.8 展示了一个 4 节点网络结构图,其中所有点之间都有直接连接。行为人 A 与 B、C、D 都有直接联系,测地距均为 1。因此,行为人 A 的亲密中心度值为 $1/(1+1+1)=1/3$。虽然图 4.8 中的行为人 A 与其他关系人的直接联系要比图 4.7 中的行为人 A 更多,但二者的亲密度值是一样的,因为第二个网络更大。为便于比较,标准化亲密中心度将网络规模考虑其中,对这两种情形进行了区分。图 4.7 的行为人 A 的标准化亲密中心度值为 $(3-1)(1/3)=0.67$,而图 4.8 的行为人 A 的标准化值更高一点,等于 $(4-1)(1/3)=$

1.0。行为人的亲密中心度值越高,那么它与其他行为人的关系越紧密,从某种意义上来说,就是该行为人可以通过较短的测地距通达其他所有的节点。

与组程度中心度相似,组亲密中心度也是一种离差测量,表示在一个网络内亲密中心度的等级分层。具体来讲,组亲密性中心度测量一个给定网络内行为人各自亲密中心度之间的差异程度。根据弗里曼(1979)的文献,组亲密中心度指数的计算公式如下:

$$C_C = \frac{\sum_{i=1}^{g}[C'_C(N^*) - C'_C(N_i)]}{[(g-2)(g-1)]/(2g-3)} \qquad [4.7]$$

其中 $C'_C(N^*)$ 表示网络中可观测到的行为人亲密中心度的最大值,而 $C'_C(N_i)$ 是 $g-1$ 个其他行为人的亲密中心度。当一个网络中行为人的亲密中心度分布极其不均时,组亲密中心度达到最大值,为 1.0。此时,某一行为人具有最大的亲密中心度,而其他所有人只有最低的亲密中心度。相反,如果每一个行为人都有相同的亲密中心度,那组亲密中心度等于 0。

中间性中心度

中心度的中间性这一概念关注其他行为人如何控制或者调整并不直接连接的两节点间的关系。行为人中间性中心度测算其他行为人在何种程度上依赖网络中两节点间的测地距(即最短距离)。中间性中心度是一个非常重要的在一个网络控制信息交流或资源流动的指标。假设行为人 j

必须通过行为人 i 才能与行为人 k 沟通，因此行为人 i 负责或控制行为人 j 与 k 之间所传递信息的内容及时间。行为人 i 在众多对偶关系点间的测地距路径中出现的次数越多，行为人 i 控制网络交流的可能性就越高。

为了量化行为人 i 的中间性中心度，弗里曼(1977)提出下列处理步骤。首先，g_{jk} 是点 j 和 k 间测地距的路径数量，$g_{jk}(N_i)$ 是点 j 和 k 之间所有经过点 i 的测地距路径的数量。然后，将 $g_{jk}(N_i)$ 除以 g_{jk}，便得到连接点 j 和 k 的所有经过点 i 的测地距路径的比例。将所有不包括点 i 的对点加起来，便是测算节点 i 在其他网络成员间测地距路径的位置程度。下面的公式反映了这个逻辑：

$$C_B(N_i) = \sum_{j<k} \frac{g_{jk}(N_i)}{g_{jk}} \qquad [4.8]$$

当 $g-1$ 个节点间所有对点的测地路径都不经过点 i 时，该中间性中心度值为 0.0。假设每一对点只有一条测地路径，当节点 i 经由所有对点的每一条测地路径时，该值即达到最大值 $(g-1)(g-2)/2$ $\left(\text{除节点 } i \text{ 之外，} g-1 \text{ 个节点之间的所有测地距路径数量是 } C_{g-1}^2 = \frac{(g-1)!}{2!(g-1-2)!} = \frac{(g-1)!}{2!(g-3)!} = \frac{(g-1)(g-2)}{2}\right)$。然而，假如每对点有不止一条测地距（最短路径）时，节点 i 的中间性中心度的最大可能值将超过 $\frac{(g-1)(g-2)}{2}$，具体值依每对点间的测地距路径的数量而定。

沃瑟曼和浮士德（1994：190）通过 $C_B(n_i)$ 除以

$\frac{(g-1)(g-2)}{2}$ 的理论极大值来得到标准化行为人中间性中心度 $C'_B(N_i)$（当然，还是假定每对节点间只有一条测地距路径），其公式为：

$$C'_B(N_i)=\frac{C_B(N_i)\times 2}{(g-1)(g-2)} \qquad [4.9]$$

当初始中间性中心度是 0 时，标准化行为人中间性中心度值也是 0.0；当其他 $g-1$ 个节点中每对节点间的测地距路径都含节点 i 时，标准化行为人中间性中心度是 1.0。因此，该值越接近 1.0，节点 i 就越能控制或调解网络关系。

如同组程度和亲密中心度测量一样，组中间性中心度测量该值在网络中不同行为人间的变化程度。按照弗里曼(1979)的基本测算方法，沃瑟曼和浮士德(1994)提出了组中间性中心度指数：

$$C_B=\frac{\sum_{i=1}^{g}[C_B(N^*)-C_B(N_i)]}{((g-1)^2(g-2))/2} \qquad [4.10]$$

分子是最高值行为人和每个其他行为人的中间性中心度值的差异总和。分母表示网络中所有节点的中间性中心度理论上的最大可能值。行为人中间性中心度在 $\frac{(g-1)(g-2)}{2}$ 时获得理论最大值。在群体层次，如果某一主要节点居间促成了所有其他的测地线，那这个最大值至少发生 $g-1$ 次。因此，对于一个含有 g 个行为人的网络，其理论最大中间性中心度值是 $\frac{(g-1)^2(g-2)}{2}$。（这里我们再次强调，这个计算假定每一对点只有一条测地距路径。如果两

点之间存在多条测地距路径，那么行为人中间性中心度的最大可能值将大于 $\frac{(g-1)(g-2)}{2}$，这将导致理论最大可能值的相应变化。）

当所有的测地线都经由某一主要节点时，组间中心度为 1.0。相反，如果在一网络中每一个节点都有相同的中间性中心度，那么分子为 0，组层次中心度也为 0.0。因此，中间性中心度越接近 1.0，网络内中间性中心度越是不均衡分布。

声望

许多情况下，社会网络数据收集包括指导如何在调查中确定网络关系的发起人或发送人和终结者或接收者。在这样的有向图中，仅仅参与或融入某一关系，要比区分这一关系的发起者和接收者次要得多。例如，在一个工作场所的报告关系网络中，低阶位的员工通常向他们的管理监督者报告他们的工作行为和绩效工资绩点，而高层雇员很少或根本就不用对他们的下属报告工作。在一个朋友关系网络中，一个热情地将许多人视为好朋友的人，可能没有人将他视为好友。

我们将声望定义为一个网络中的一个社会行为人接收他人所发出的关系或被他人视为关系客体的程度。这种发送者—接收者或源头—目标的区别，极大地强调了对资源控制上的不平等，以及伴随这种不平等而来的权威和服从(Knoke & Burt, 1983:199)。

根据这个定义，声望可以简单地通过计算某一个行为人接收其他行为人某种特定关系的有向连接数量来测量。例如，罗曼和诺克(1987:163)询问美国能源和健康政策领域的知情人士，请他们确认“对制定国家能源/健康政策具有特别的影响力和重要性”的机构名单。每一个提名都得到平等对待，没有将其来源进行区别对待，这样所获提名便反映了一个行为人在系统内的受欢迎程度。沃瑟曼和浮士德(1994:202)指出，行为人程度声望可以根据有向图的节点入度来计算：

$$P_D(N_i) = \sum_{j=1}^{g} x_{ji} \, (j \neq i) \qquad [4.11]$$

因为这一测量加总由 g 个节点所组成的非对称社会矩阵中的第 i 列的值，因此它的最小值是 0.0，最大值是 $g-1$。一个标准化的行为人程度声望，是控制了网络规模的大小，可以对不同网络进行声望比较：

$$P'_D(N_i) = \frac{\sum_{j=1}^{g} x_{ji}}{g-1} \, (j \neq i) \qquad [4.12]$$

当所有其他行为人都提名行为人 i 时，标准声望值为 1.0；当无人提名行为人 i 时，标准声望值为 0.0。因此，这一比例越接近 1.0，行为人的声望越高。

上述段落中所介绍的中心度、组中心度和声望的基本测量方法，都是基于二进制数据的。除了弗里曼、博加蒂和怀特(Freeman, Borgatti & White, 1991)讨论的赋值图中的中心度问题之外，对赋值图的相关测量研究极少。很显然，对于赋值图的中心度和声望测量问题，需要更多的研究。

自我中心网络中心度

中心度和声望的分析局限于全幅网络数据，直到 Marsden(2002)将中心度的测量延伸至自我中心网络数据。他讨论了一个单一关系的二进制对称数据的中心度测量。在使用调查对象进行的自我中心网络研究中，关系主体提名 g 个与他有特定关系的客体。每一个关系主体 i 产生一个 $(g+1)\times(g+1)$ 的数据矩阵 A_i，它的行和列都由关系客体 1 至 g 加上主体 i 构成。因为在定义中，主体与所有其他的客体具有直接联系，因此第 i 行中的每一个单元是 $x_{i,j}=1$，其中 $1\leqslant j\leqslant g$，而且 $i\neq j$。

在一个全幅网络中，行为人程度中心度是行为人与他人的直接联系(参见公式 4.1)。在含有 g 个关系客体的自我中心网络中，关系主体 i 的程度中心度是行为人程度中心度的最大可能值 $g-1$。关系主体 i 的标准化程度中心度总是 $\frac{g-1}{g-1}=1.0$。全幅网络中的行为人亲密中心度，是行为人 i 与其他行为人之间测地距总和的倒数(参见公式 4.5)。标准化亲密中心度控制了网络规模(参见公式 4.6)。因此，行为主体的亲密中心度和标准化亲密中心度分别为 $\frac{1}{g-1}$ 和 $\frac{g-1}{g-1}=1.0$。最后，全幅网络中的中间性中心度测量节点 i 经过所有对点间测地距的位置程度(参见公式 4.8)。由于行为主体 i 在其自身的自我中心网络中与所有的行为客体直接连接，因此 i 可以作为每对并不直接相连的客体间的中介点。

马斯登(2002:410)宣称,在自我中心网络中,节点 i 的中间性中心度的测量不同于全幅网络。一方面,在自我中心网络中,节点 i 的中间性中心度会偏小。自我中心中间性测量不包括主体节点 i 的中间位置,这中间位置是任何一对经由超过三个路径长或更长的测地距的节点之间的位置,但是在全幅网络的中间性测量中则包括这种情形。另一方面,当两个客体节点的连接不但可以经由节点 i 连接,而且还经由节点 i 所处自我中心网络之外的其他节点时,自我中心网络中间性中心度会夸大主体节点 i 的中间性中心度。在这种情况下,自我中心网络中间性测量会忽略此网络之外的中间节点,而全幅网络中间性中心度测量则会包括这种中间节点。

关于这些观点的描述,请参见图 4.9。行为主体 i 的自我中心网络由在某种关系下提名的关系客体 A、B、C、D 组成。全幅网络也包括节点 M,它连接节点 i 的两个客体(A 和 D),虽然 M 本身并不是节点 i 自我中心网络的一部分。节点 i 的四个关系客体构成六对关系(AB、AC、AD、BC、BD、CD),其中每一对都有一条独特的测地距路径。由于只有 BC 是直接相连的,而其余每对关系的测地距路径都包含主体点 i,因此 i 的中间性中心度在自我中心网络中是5/6。然而节点 M 的出现,使得节点 i 的自我中心网络中间性中心度

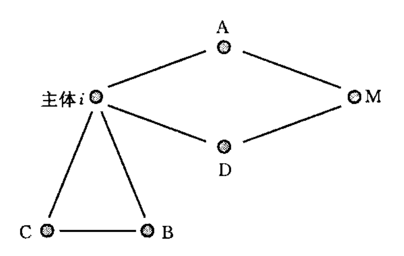


图4.9 自我中心网络的中间性中心度

和全幅网络中间性中心度之间产生了差别。一方面，节点 i 自我中心网络中间性测量偏小，这是由于忽略了节点 i 在两条测地距（对点 MB 和 MC）中的媒介位置。另一方面，由于忽略了 AD 间的测地距通过点 M 的情形（这与通过节点 i 的 AD 间测地距形成了竞争），节点 i 的自我中心网络中间性值又会偏大。

尽管在自我中心网络和全幅网络中间性的测量中存在这些差异，但马斯登（2002）通过分析 17 个实证网络发现这两种测量方法相差不大。因此，当全幅网络中相关数据难以获得时，自我中心网络中间性中心度似乎可以作为全幅网络的行为人中间性中心度的可靠替代。马斯登（2002）认为，自我中心网络和全幅网络中间性结果的差异也许是由于两种网络数据收集的不同造成的，而这是一个值得进一步系统性研究的重要课题。

第5节 | 小团体

内聚力与一个社会群体或封闭社会关系圈中的亲密程度的概念密切相关。内聚型亚群体由通过多种直接互惠的选择关系相连的行为人构成。他们之间能够共享信息,团结互助,一致行动。所有的亚群体成员间有多种直接联系,很少或几乎不包含外来者,倾向于构成一个在思想、认同和行为方面具有同质性的群体。内聚型亚群体的例子包括异教组织、恐怖活动小组、犯罪帮派、军事行动小组、运动团队、手工行业和工作团队等。“小团体”已经进入了日常语言之中,主要指学校、教会和俱乐部这样的高层群体。这个术语在社会网络分析中具有特定含义,源于诸如节点、线和路径长度这些图论要素。利用小团体的概念来分析研究群体结构,可以帮助研究人员更好地理解内聚力是如何通过提供建议和利益支持来帮助群体成员的(Dunbar & Spoor, 1995),更好地理解对小团体的广泛信赖是如何限制网络联系的范围的(Blau, Ruan & Ardelt, 1991)。

社会学的概念,诸如组、群、圈子、帮派、派系、小团体等并没有严格的区别,可以互换(Borgatti, Everett & Shirey, 1990)。在对社会网络研究中大量关于亚群体的文献进行总结后,沃瑟曼和浮士德(1994:251)分析出了关于内聚型亚群

体的四个普遍特征:连接的相互性、亚群体成员间的可联络性、成员间联系的频率,以及与圈内成员间相对于与非圈内成员间联系的相对频率。这些特征是定义小团体和网络亚群体相关测量的基础。

小团体

一个小团体是含有3个或多个节点的最大完全子图,其中所有的节点都互相直接连接,网络中的其他点都没有与小团体中的每一个成员有直接联系。因此,在一个小团体中,每对节点间的测地距都是1.0。也就是说,每两个节点间都由一条直接路径相连。图4.10描述了一个含有6个节点的二进制无向性网络中的小团体。两个小团体ABE和BCDE,每一个都符合必要条件。虽然F与A和C都有连接,但这3个点并没有构成一个小团体,因为A和C没有直接联系。4个其他的小群体(EBC, EBD, BCD, BDE)也不是小团体,因为另外还有其他的点与这些小组别的所有三个点直接连接。例如,EBC并不是一个小团体,因为点D与E、B、C直接相连。只有BCDE这个最大的完全子图符合小团体的条件。

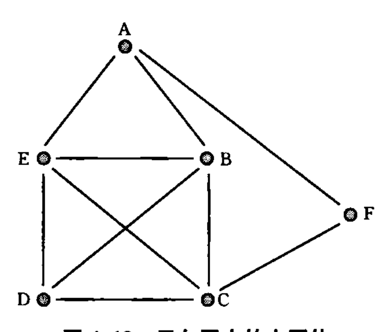


图4.10 无向图中的小团体

作为识别内聚型亚群体的一个方法,小团体确认规则有着非常严格的条件。小团体概念有一个“吝啬鬼”的名声(Alba, 1973),这是因为仅仅某一条线的缺失也许就会让一个子图不能成为一个小团体。因为小团体构成的高门槛,实证研究者极少在低密度的实际网络中发现大的小团体(Wasserman & Faust, 1994:256)。在实际网络数据中很少发现小团体的原因之一在于网络数据收集的设计。例如,固定名单法可能会限制受访者提名的关系客体的最大数量(如“说出你的三个最好的朋友”)。因此,小团体的规模不可能超过研究人员设定的上限。

小团体标准也严格地将内聚型亚群体内成员与群体外成员区别开来,因此忽略了更核心的和更边缘的行为人之间的等级层次。一旦小团体的边界划定了,在小团体和非小团体成员间的区分就很难更为精细了。如此简单的一分为二法则常常是不够的,因为它不能反映更为重要而微妙的差别。一个图中小团体数量的决定因素之一是全幅网络的规模。小规模网络极少形成小团体,而大规模数据库常常产生大量的微小团体,而且它们还常常互相重叠。这一限制导致研究者集中在多重重叠的小团体,而不去搜寻一个单一的独立小团体。

n-小团体

虽然严格的小团体定义使得它拥有的信息量有限,但是它紧接着也催生了有关亚群体内聚力的测量,从而在一定程度上缓解了小团体构成方面的严格要求带来的压力。一个

通常的方法就是基于节点度数，即通过某点的线条的数量(Doreian & Woodard，1994；Seidman，1983)。我们在第3章论述k中心概念时讨论过这个方法。这一方法将k中心定义为一个亚群体，条件是每一个群体成员至少与k个其他点直接连接。通过改变k值，研究者能够为定义一个亚群体的边界设定一个严格或宽松的标准。

第二个通常的方法是利用节点连接来修正小团体(Alba，1973；Mokken，1979)。n-小团体概念放宽了小团体本身的严格条件，即每对节点间的测地距值必须是1.0。在n-小团体中，所有对点间的最大测地距是路径长度n；也就是说，没有节点离其他每个节点的距离超过n个连接。通过改变n的值，网络研究者可以区分具有更强凝聚力的亚群体(n小)或者更弱凝聚力的亚群体(n大)。例如，设定n值等于2，即等于一个作为有3节点或多节点的子集的2-小团体，其所有的测地距是2或更小；而网络的其他所有节点都与2-小团体连接，其测地距是3或大于3。在无向图(图4.10)中，节点F在严格标准下不属于任何一个小团体。

然而，将测地距从1放宽至2，即可得到一个包含节点F的2-小团体(整个图为ABCDEF)，因为该节点(同时节点A)连接2-小团体的所有成员，而其中的测地距等于或小于2。n-小团体标准具体化为：任何对点间的最大测地距不能超过n。一般来说，n的值越大，n-小团体的包容性也就越强，因此其成员间的凝聚力也就越弱，因为一些对象之间距离要达到n步。严格定义的小团体是n-小团体的一个特例，其n值等于1.0。

有向图中的小团体

有向图将关系的发送者和接收者区别开来。在有向图中识别小团体需要注意对偶连接关系的互惠性。为拓展无向图中小团体的严格定义,我们将有向图小团体定义为由等于或大于3个节点构成的子图,其中所有的每一个节点都直接与其他节点连接,而且每对点之间都是对接的(即对偶连接)。为了描述有向图的小团体,图4.11展示了一个由6个节点组成的有向连接网络。在这个网络中存在三个3节点小团体,分别是ABE、BDE、BCD。由于从节点E到C的关系是非互惠的,因此EBCD不构成一个有向小团体。同样,A和C之间由于对偶关系的缺失也导致AFC不能成为一个有向小团体。

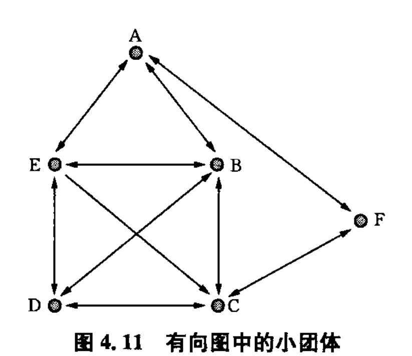


图4.11 有向图中的小团体

从无向图 n-小团体的定义得到启发,研究人员提出,通过改变对偶连接的对点间的测地距值,可以放宽小团体定义上的严格要求。这一方法与前面的方法具有极大的相似性,除了更多的考虑有向连接问题。例如,2节点可以通过四种

情形生成 n-连接:弱 n-连接、单向 n-连接、强 n-连接、双向 n-连接(Peay, 1980; Wasserman & Faust, 1994:275)。在最强形式的 n-连接——双向 n-连接——中,从 i 到 j 的最短路径和从 j 到 i 的最短路径必须通过同样的节点,不过是方向相反而已,而且其路径长度等于或小于 n。这四种情形的每一种都构成了一个有向 n-小团体,产生了四种类型的 n-小团体亚群体。在图 4.11 中,子图 ABCDF 是一个双向 2-连接,因为它的 10 个对点通过测地距 1 或 2 都是互相可及的。然而,E 不包括在内,因为它得不到 C 的反向连接,导致由 E 至 C 的测地距长要短于由 C 至 E 的两条测地距长(C-D-E 和 C-B-E)。

第6节 | 结构等价

社会科学家常常不仅对行为人的内聚力感兴趣，还对行为人的平等状态感兴趣。这种平等指在一个网络中两个或多个行为人与他人具有同样的或非常相似的关系。结构等价的行为人之间一般来说具有竞争型而非内聚型关系。譬如，两个大白菜种植者都将他们的产品卖给同一批蔬菜批发商，因此在出售大白菜时他们的关系就是结构等价，处于激烈竞争状态。结构等价的行为人完全可以互相替代。如果一个种植者退出蔬菜供应网络，那么他的位置很容易就被另外的结构等价的种植者替代，从而保持原来的网络不发生变化。为了从其他网络参与者那里获取适宜的反馈，社会网络中的完全可替代性常常导致激烈的竞争，众所周知的例子就是小学生为了获得老师的特别青睐而互相竞争。利用结构等价方法的网络研究学者往往对理解竞争关系有兴趣，而不是对群体内聚力有兴趣(Burt, 1992)。

与小团体的定义类似，结构等价的定义也是非常严格的。在有向二进制图中，如果两个行为人在与他人发送和接收关系时具有完全相同的连接模式，那么他们在那种具体关系上就是完全结构等价的。更详细地讲，对于网络中所有的节点 k(但不包括节点 i 和 j)，如果节点 i 发送一个连接给 k，

当且仅当节点 j 也发送一个连接给 k；如果节点 i 从 k 接收一个连接，当且仅当节点 j 也从 k 接收一个连接，那么节点 i 和 j 就是结构等价的（Wasserman & Faust，1994：356）。对于多重网络来讲，如果两点结构等价，则上面讲的这个条件必须在 R 种关系的每一个当中都要精确把握。节点 i 和 j 之间有向连接的存在与否，对于决定它们是否结构等价是无关的。它们的结构等价仅仅是由它们与 $g-2$ 个其他网络点的关系模式来决定。

无向二进制图不区别关系的发送者和接收者。因此，将上述有向关系结构等价的定义沿用至无向图，如果对于网络中所有的其他行为人 k，i 与 k 具有连接，当且仅当点 j 也与 k 具有连接，那么 i 和 j 就完全是结构等价的。结构等价也可以用在赋值图中，其中利用排名等级而非二进制赋值来测量两节点之间的连接。严格来讲，只有当赋值图的两个节点在与其他节点的每一个连接中都具有完全同样的值，这两个节点才是结构等价的。

对于实证网络分析的实际运用来说，完全结构等价性的上述定义通常太过严格。真实网络数据中极少有数据能够满足这样苛刻的标准。然而，当一些节点与其他节点的关系模式高度相似，尽管不是完全相同，但我们也可以认为这些节点是大致结构等价的。为了获得这种近似性，研究人员利用关系相似性的测量，而不采用严格的、非此即彼的完全结构等价的标准要求。两个节点各自与其他所有节点的连接越相似，那这两个节点就越接近结构等价。

基于网络关系相似性的两节点间的结构等价，只要求他们与其他节点或有或无的关系模型高度相似即可。假定存

在一个二进制有向图，两个结构等价的行为人，在其社会矩阵对应的行和列中的条目非常雷同。根据这一原则，伯特(1978)提出将欧氏距离作为测量行为人 i 和行为人 j 结构等价的一种方法：

$$d_{ij} = \sqrt{\sum_{k=1}^{g}[(x_{ik} - x_{jk})^2 + (x_{ki} - x_{kj})^2]}(i \neq j \neq k)$$

[4.13]

其中 d_{ij} 是行为人 i 和行为人 j 之间的欧氏距离，x's 是社会矩阵(第一个下标表示行，第二个下标表示列)的值(对于二进制关系来说就是 1 或 0)。由于 d_{ij} 是两个差值的平方和的正平方根，因此每个 $d_{ij} \geqslant 0$。假如行为人 i 和行为人 j 与其他行为人之间拥有非常一致的连接，那么所有的差值 $(x_{ik} - x_{jk})$ 和 $(x_{ki} - x_{kj})$ 都是 0；因此这两个完全结构等价行为人的欧氏距离 $d_{ij} = 0$。但是在绝大多数实际例子中，观察到的 d_{ij} 值要大于 0；因此，行为人 i 和行为人 j 在某种变动的程度上关系类似。欧氏距离与行为人的相似性相反，因此也与结构等价相反：d_{ij} 值越大，行为人 i 和行为人 j 的结构等价度就越小。也就是说，欧氏距离实际上测量的是一对行为人之间的异质性。

为了阐述如何计算对点之间的欧氏距离，图 4.12 和表 4.3 分别描述了一个 5-节点网络结构的有向图和矩阵式。图 4.12 表明，行为人 A 和行为人 B 是完全结构等价，因为他们都与行为人 C 和行为人 D 有直接联系，但都与行为人 E 没有联系。相反，行为人 D 和行为人 E 就不是结构等价了，因为尽管他们都发送了一个连接给行为人 C，但是行为人 D 接收了行为人 A 和行为人 B 的连接，而行为人 E 却没有得到任何连接。利用表 4.3 的二进制值，可以计算行为人 A 和行为

人B的欧氏距离：

$$d_{AB}=\sqrt{\begin{array}{l}[(x_{AC}-x_{BC})^2+(x_{CA}-x_{CB})^2]+[(x_{AD}-x_{BD})^2+\\(x_{DA}-x_{DB})^2]+[(x_{AE}-x_{BE})^2+(x_{EA}-x_{EB})^2]\end{array}}$$

$$d_{12}=\sqrt{\begin{array}{l}[(1-1)^2+(0-0)^2]+[(1-1)^2+(0-0)^2]+\\[(0-0)^2+(0-0)^2]\end{array}}$$

[4.14]

其中 $d_{12}=0$，表明行为人A和行为人B是完全结构等价。你能否算出点D和E之间的欧氏距离是 $\sqrt{2}=1.41$？

由于无向二进制图不区分关系的发送者和接收者，因此它的欧氏距离计算就简单得多：

$$d_{ij}=\sqrt{\sum_{k=1}^{g}(x_{ik}-x_{jk})^2}(i\neq j\neq k) \qquad [4.15]$$

当网络呈现多重关系时，欧氏距离就得计算所有 R 种关系的对点差值的平方和：

$$d_{ij}=\sqrt{\sum_{r=1}^{R}\sum_{k=1}^{g}[(x_{ikr}-x_{jkr})^2+(x_{kir}-x_{kjr})^2]}(i\neq j\neq k)$$

[4.16]

与通过欧氏距离来计算异质性不同的是，皮尔逊相关系数 r_{ij} 直接测量关系的相似性，其中对点关联值越高，表明结构等价度越大。我们将在本节块模型部分讨论相关性问题。

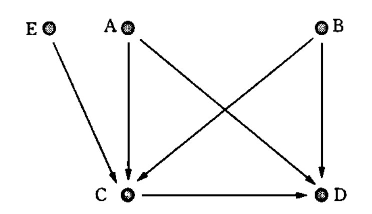


图4.12 5-节点网络的结构等价有向图

表 4.3 5-节点网络的结构等价矩阵

	A	B	C	D	E
A	—	0	1	1	0
B	0	—	1	1	0
C	0	0	—	0	0
D	0	0	1	—	0
E	0	0	1	0	—

第7节 | 视觉演示

网络图像在社会网络研究中普遍用于展现和传递网络结构的亮点。社会网络分析在视觉演示的发展过程中经历了三个明显的阶段(Freeman, 2000, 2005)。起始阶段大概在20世纪30年代,当时雅克布·莫雷诺发明了一种手绘的即兴社会网络图,用来描绘诸如学生等行为人之间的关系(Moreno, 1934)。这种自由式的方法很快就让位于通过标准计算程式来标示点和线的图。可视化初始阶段的一个基本原则就是所示空间应该代表行为连接的模式,即在数据矩阵中社会关系最近的对点,在可视图像中也能看得出来是距离最近的(Freeman, 2005)。然而,在二维或三维的视觉展示图中,精确标注多位行为人之间关系的亲近度和距离,通常是不可能完成的任务。因此,研究者最终采用系统简化、维度减少但依然能够反映原始数据模型的办法。可视化第二阶段始于大约20世纪60年代,用大型电脑主机和软件自动生成图表。尤其需要指出的是,网络分析不断利用这部分将要讲到的分层聚类法和多维标度法。最近的阶段始于20世纪90代中期,随着因特网、高速计算机网络和浏览器,以及个人电脑普及化时代的到来,关系数据库的大型可视化展示,包括纵向网络变化的动画展示,就有了新的机会(如

Brandes, Kenis, Raab, Schneider & Wagner, 1999; de Nooy, Mrvar & Batagelj, 2005)。

因为篇幅有限,我们就集中讨论分层聚类和多维标度这两个展示网络数据的基本理念。可视化展示对于用社会网络数据来探索发现内聚型亚群体以及揭示它们内部的关联性是非常有用的。在此过程中,可通过将众多行为人概括为数量较少的共同占据位置,对复杂的网络结构进行简化。结构等价方法计算和分析对点异质性(欧氏距离)或相似性(关联指数)的矩阵模型,来确认由行为人与其他人相同或相近关系模型所构建的亚群体所占据的网络位置。聚类和多维度标示法运用于结构等价矩阵,来确认共同占据位置,以二维或三维图的形式来展示结果。

聚类

分层聚类分析,或者简称为聚类分析,将行为人分成亚群体(共同占据位置)。这些亚群体成员是标准的或近似的结构等价。每一个行为人开始都被当成一个单独的聚类,然后这些单独的聚类依次联合起来,直到合并成为一个独立的聚类。一个树图,也被称为树形图,形象地展示了这个合并聚类的层次顺序。

聚类演绎法　一般处理 $g \times g$ 阶方阵的欧氏距离(d_{ij})或关联指数(r_{ij}),其中 g 是具体关系中的行为人数量。聚类分析的一项比较研究发现,这两种方法都得到非常相似的结果,虽然使用相关性概念的聚类分析会相对容易解释(Aldenderfer & Blashfield, 1984:24—28)。多重网络关系

(R)可以同时被聚类，通过构建一个 $g \times g \times R$ 维的数据阵列，在进行聚类分析前，计算所有 R 矩阵中的每对点间的欧氏距离或相关性。虽然分层聚类可用于二进制数据和赋值数据，但为简化的缘故，我们主要对欧氏距离的二进制对称矩阵进行讨论。

在计算对点的欧氏距离矩阵之后，聚类分析开始根据一个阈值 α 对行为人进行合并。这个阈值作为分析者的一个上限，让他们决定在一个结构等价的特定水平上，哪一个行为人属于同一个关联性占据位置(聚类)。仅当 $d_{ij} \leqslant \alpha$ 时，行为人 i 和 j 共同占据一个位置。在一个聚类中的行为人互相之间的社会距离要小于他们与其他聚类中行为人的距离，即聚类内的行为人更具有结构等价性。通过聚类演绎法，不断将 α 的值放宽(即 α 值变大)，并运用于整合行为人的过程，直到整个网络合并成一个单一的全涵盖聚类。尽管分层聚类法获得一些非交叉聚类，但聚类还是嵌套的；也就是说，较小的聚类依次包含在 α 值高的较大聚类中(即，结构等价度越低，相似性越小，那么聚类内距离就越大)。最终，研究人员必须决定哪一个层面的块(即哪一个 α 值)能够为网络中结构等价位置数量提供最好的替代。

研究者可从形成聚类的三个基本标准中进行选择：单个联结、平均联结或完全联结。给定 α 值的水平，当两个关系最为亲近的行为人距离小于 α 时，单个联结标准将两个聚类合并成一个聚类。在完全联结标准下，当每一对行为人的距离小于 α 时，两个聚类进行合并。平均联结标准是一个折中的选择，要求两个合并的聚类中两类行为人的平均距离小于 α。实证研究发现，每一种选择都存在有利和不利的一面，研

究者应当根据具体的情况来调整所选的聚类分析方法(Aldenderfer & Blashfield, 1984:53—62)。完全联结聚类倾向于得到数量众多的同质型和紧凑型的聚类,而产生"链"的可能性较低。"链"在这里指依次增加一个行为人即可形成一个单一的较大的聚类(Burgin, 1995; Wasserman & Faust, 1994:381)。

前面的内容主要讨论了小团体、聚类和结构等价等有关的概念和方法。这些网络分析方法普遍用于揭示社会结构的基本要素,诸如社会内聚力(小团体和聚类)或结构性竞争的程度(结构等价性)。因为这些方法探察社会结构的中间层次,它们承上启下,将社会网络的微观模式与社会结构的全景网络连接起来。下面我们讨论多维标度,这是揭示网络行为人结构的另一个非常重要的方法。

多维标度

多维标度(MDS)是将数据结构进行可视化处理的第二个基本方法。MDS 的主要目的是,发现那些能够反映网络行为人之间相似性和差异性(距离)的有意义的维度。如同聚类分析所述,MDS 的典型要素是一个欧氏距离为(d_{ij})或关联指数为(r_{ij})的 $g \times g$ 阶矩阵,其中 g 是具体关系中行为人的数量。MDS 的可展现结果是一张图,或称为社会关系图,其中较短距离(较大相似性)的关系行为人比较远距离(较大差异性)的关系行为人在空间上要离得近一点。虽然 MDS 图能够展示多个维度,但大多数分析一般都只是表达二至三个维度。

在MDS图中所表示的行为人之间的距离与输入矩阵图中的相似性或差异性值相关但不完全相同。这个距离反映的是通过MDS程序所估算到的这些数据中对点间的距离。应力这一概念测量所观测到的矩阵和所计算到的矩阵二者之间所有点的差异性(Kruskal & Wish, 1978:23—30):

$$Stress = \sqrt{\frac{\sum\sum(f(x_{ij}) - d_{ij})^2}{Scale}} \qquad [4.17]$$

其中,$f(x_{ij})$是输入值的一个非计量性的、单一的函数(Kruskal & Wish, 1978:29),d_{ij}是呈现在图中行为人i和行为人j之间的欧氏距离。Scale是一个应力值在0.0至1.0之间的测量指示因子。如果一个MDS图将一组输入数据转化成标准的图值,那么对于所有的i和j来说$f(x_{ij} = d_{ij})$,应力将等于0。所以,应力值越低,MDS图的空间就越能够代表所观测到的网络行为人的社会距离。基于这个原因,应力概念也被称为MDS拟合度的测量法,其中应力值低于0.1被认为是拟合得非常好,应力值在0.1至0.2之间被认为是拟合度适中,应力值大于0.2就拟合较差(Kruskal & Wish, 1978:52; Slez & Martin, 2007)。

MDS曾经被运用于各种研究。例如,斯莱兹和马丁(Slez & Martin, 2007)利用MDS对各州一段时期内随时间变化投票的相似性进行绘图比较。维特、金和格利克(Whit, Kim & Glick, 2005)对多伦多地区的50个种族群体绘制了社会空间位置图。MDS的最大优势,是允许他们同时分析所有的种族群体,而传统的分隔研究方法则将每一个群体与所有其他群体进行分离对比。

为了描述聚类和 MDS 这两种方法是如何能够一起联手揭示结构等价网络位置,我们可以看图 4.13 和图 4.14。图 4.13 分层聚类树图和图 4.14 MDS 图都是基于欧氏距离,而该欧氏距离是从图 4.12 的二进制网络数据计算出来的。聚类分析使用完全联结选择:当所有行为人之间的距离小于 α 值时(这些值显示在图 4.13 图的水平轴上),聚类合并在一起。例如,行为人 B 和 E 间的欧氏距离(α)是 0.00(标准的结构等价),因此树图首先在欧氏距离等于 0.00 这一阶段将 B 和 E 合并。有两个聚类出现在欧氏距离等于 1.00 这一阶段:一个聚类合并了行为人 A 和 C,另一个聚类将 D 与 B—E 这一现存的聚类合在一起。在欧氏距离 $\alpha = 1.73$ 时,这两个聚类合并了,表明所有对点的欧氏距离都小于 1.73。最后,在欧氏距离 $\alpha = 200$ 时,行为人 F 与 A—B—C—D—E 聚类合并,表明 F 和所有其他节点之间的欧氏距离在 1.73 至 2.00 之间。

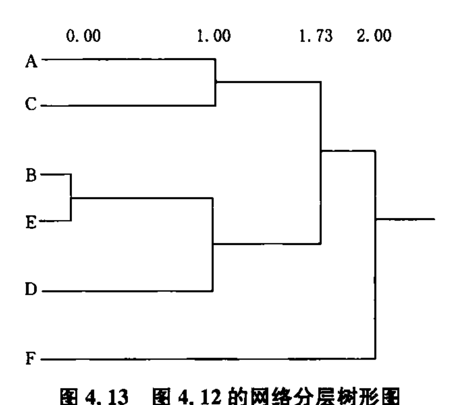


图 4.13 图 4.12 的网络分层树形图

图 4.14 的 MDS 图中,行为人亚群体周边有邻接线,这个位置正好是在聚类树图中 $\alpha = 1.00$ 所确认的三个聚类所占的位置。

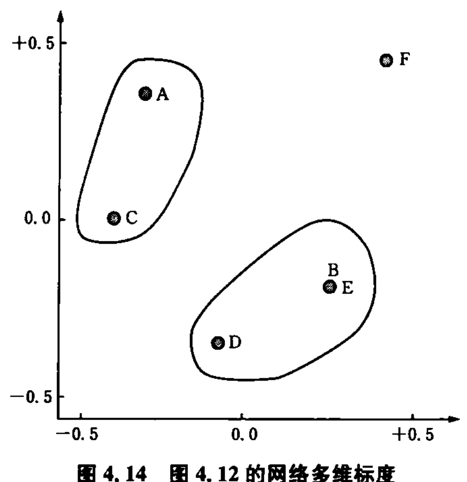


图 4.14 图 4.12 的网络多维标度

行为人 B 和 E 在图中正好占据着同样的点，因为他们是完全结构等价。虽然对于确认聚类来说，α 值的选择在一定意义上是随意的，但我们仍然建议选择一个值，它能够平衡网络行为人之间的内聚力和分裂性。例如，如果我们已经知道欧氏距离为 0.00，那么我们将有五个聚类：B 和 E 构成一个聚类，另外四个行为人占据单件聚类的位置。为方便比较，现假设我们选择 $\alpha = 2.00$，那么结果只有两个聚类：(A—B—C—D—E)和(F)。第一个选择标准，由于其亲密关系设置上的高门槛，产生了大量的分裂性；而第二个值，则使得门槛降低，得到了较大的内聚力。然而，没有一个方法能够像中间值 $\alpha = 100$ 时提供的信息，即揭示联合占据位置之间的结构关系，为这个网络的分离性与内聚力提供一个更平衡的视角。

MDS 的一个延伸运用是加权多维标度(WMDS)。它在传统的 MDS 方法中增加了加权因素，用于代表矩阵间变化

的信息。因此,WMDS允许多重矩阵被进行系统性地区分,从而对距离模型进行概括。例如,假如每一个矩阵都对应一个不同的个人,那么WMDS利用权重就可以对这些个人思考和想象他们的关系进行不同的刻画。基于这个原因,WMDS也被称为个体差异标度(INDSCAL)。虽然我们不能在这里更详尽地讲述WMDS,但我们鼓励有兴趣的读者向希夫曼、雷诺兹和杨(Schiffman, Reynolds & Young, 1981: 55—85)咨询;还有韩(Han, 2003),他对此方法有最新的运用。

第8节 | 块模型

块模型是一个矩阵代数方法，将网络行为人分类成联合占据的、结构等价性的位置。块模型方法开始由哈里森·怀特和他的助手研发(Boorman & White, 1976; Schwartz, 1977; White et al., 1976)。自此开创性工作之后，研究者已经将块模型方法成功地运用到各种问题研究中，从跨组织网络(Knoke & Rogers, 1979)到新技术的扩散(Anderson & Jay, 1985)，再到世界体系中的城市角色(Alderson & Beckfield, 2004)。方法研究者也加深了块模型搜索和划分法则的研究(Nowichi & Snijders, 2001; Winship & Mandel, 1983; Wu, 1983)。因为篇幅有限，我们在这里不继续讨论这个大的话题，在本节集中讲述块模型的基本问题及其应用和结果的阐释。

块模型是 g 个行为人社会网络矩阵的一个部分，它将一个或多个关系网络，分成两个或多个被称为块(B)的亚群体或位置。块这一术语指结构等价行为人的一个亚方阵，其中的行为人与其他块中的行为人具有相似的(如果并非完全相同的话)关系。建立块模型是一个数据简化技术，通过重新对行为人进行分组，提出浓缩的综合信息，它将对网络数据进行系统性搜索，从中找出关系模型。结果就是重新排列置

换过的密度和图像矩阵,表示每一种类关系中块内部和块之间的联结方式。块模型可以被用于单个或多重的关系,有向或无向的联结,二进制或赋值的图。我们的块模型例子是基于 UCINET 的 CONCOR 运算法则,它能够同时处理二进制数据和赋值图。我们这里只讲述二进制数据矩阵,并推荐读者阅读多里安、巴塔格尔基和费力格基(Dorian, Batagelj & Ferligoj, 2005:347—360)的相关文献,以进一步了解二进制数据和赋值图的块模型分析。

可以利用理论原则事先构建一个块模型,例如将不同管理部门的雇员分类成不同的块。然而,该方法的最普遍运用是对关系数据的经验模式进行调查搜寻。建立块模型常常使用 CONDOR(Convergence of iterated Correlations),这一方法在多个社会网络分析程序中都可以找到(Schwartz, 1977)。CONDOR 可以通过关联对行、对列,或者同时通过关联矩阵的两个向量进行运算。为教学之便利,我们这里仅讨论关联列项。首先,CONDOR 计算二进制 $g \times g$ 阶社会网络矩阵图中每一对列项的皮尔逊相关系数(对于块模型 R 多重关系来说,独立的 $g \times g$ 矩阵是堆积起来的,即安排成为一个单独的二维数据阵列,其目的是输入给 CONDOR)。这些计算不包括每对点间的直接联结,因为它们的结构等价性仅仅依赖于对点与其他 $g-2$ 个网络行为人的联结关系。假如行为人 i 和 j 与其他行为人有着几乎完全相同的联系,那么他们的相关系数 r_{ij} 将等于 1.0。相反,如果两个行为人有着几乎完全相反的联系模式,那么他们的相关系数等于-1.0。实际的相关系数几乎总是在这两个极端值之间。从 CONDOR 开始得到的结果是一个有关每对行为人的 $g \times g$ 阶对

称的关联矩阵，表示每对点结构等价性的程度。第二步，其后的所有迭代步骤，都重复着前一步关联矩阵中的对列关联过程。在一定程度上，矩阵每一格的相关系数都趋向于1.0或－1.0，直到达到这两个点时迭代中止。

接下来，CONDOR将最后的关联性矩阵排列成两个同质的块。社会网络矩阵的排列包括同时对行和列的重新安排，以将这些占据同样块位置的行为人一起排在相邻的位置。CONDOR对社会关系网络的初始分区和排列总是得到两个分矩阵，它们在行为人的数量上不一定相等，在同一块内每对行为人的所有关联指数都等于1.0，但是两个块之间的关联性却是－1.0。重复这一过程，CONDOR能够将初始得到的两个块中的每一个都续分成两个，以此类推。网络研究者必须决定在何处停止这一分化过程，因此也就决定了最终所得到的块的数量。

块模型分析导致两个主要形式的结果：密度矩阵和相应的图像矩阵。密度矩阵是一个$b \times b$阶矩阵，它的单元格的值是块内和块间的密度（密度是一个比例，其计算是由重新排列所得子矩阵中所观察的联结数量除以可能的所有联结数量）。图像矩阵也是一个$b \times b$阶矩阵，通过将密度矩阵中的每一单元格的密度值编码为0或1转化而来。下列两个可替代标准可以用来决定图像值：(1)行为人之间没有联结的格(0-块)编码为0，行为人至少有1个联结的格(1-块)编码为1；(2)研究者选择一个密度临界值α，那么在这个临界值之下的所有密度编码为0，所有等于或高于α值的编码为1。上面第一个选项是一个不现实的标准，因为实际密度为0.0的情形很少发生，除非一个网络中的关系很少，从而几乎所有图

像矩阵中的每个单元格都有 1 构成。第二个选择是最经常碰到的情形,它利用 α 密度临界值对图像值进行二项分类。研究者都选择满矩阵的密度值作为临界值。然而,因为选择 α 值不可避免地要受到研究人员的判断问题,因此一个选定的特定值容易受到独断性方面的批评。为了回应此类批评,研究人员应当尽量在理论和实践上提出合理理由,而不能仅仅诉诸利用之方便(Scott, 1991:136)。

作为块模型运用的实际范例,我们分析了美国 1988 年里根政府执掌行政权力时期国家劳工政策领域中的政策交流网络(Knoke et al., 1996)。这个政策领域的核心由 24 家最有声望和影响力的政治组织组成,由知情人士从 117 个该领域的组织中选出(参见表 4.4,组织名称首字母缩写)。图 4.15 是得到确认的有关劳工政策事项的交流网络的有向图(对于一个存在的有向联结,一对组织必须认同所发出的信息也得到了接收)。在这个值(0.29)之内的高密度联结,使得这个图中任何简单的结构都难以辨认。对应的交流关系矩阵,在表 4.4 中被转化成一个同时利用行(发送者)和列(接收者)的关联矩阵。这些关联测量每对组织与其他 22 家核心领域成员的交流模式的相似性。之后,对这个关联性矩阵进行 CONDOR 分析,得到表 4.6 这样一个含有四个模块的区分、密度和图像矩阵。

为了决定图像中的 1-模块,选定的密度临界值(α)等于或高于0.33。最后,在图 4.16中关联性矩阵的 MDS 分析,利用围绕四个模块成员的邻接线,展示了核心组织之间的距离。

在社会关系图中,联合占据的块的构成及它们的定位,反映了 20 世纪后期美国的劳工和资本之间的基本政治裂

度。在图4.16的右下角,模块1是一个由四个工会组织组成的支持劳工派,包括AFL-CIO[①],代表老年人利益的集团(AARP)和参议院劳工委员会的民主党派。较小的模块2包括三个利益集团和ETA(联邦劳工部的一个下属机构)。模块3站在支持商业的立场,由里根白宫政府、劳工部、三个商业组织,还有卡车司机工会(它不是AFL-CIO成员)。模块4也主要是站在支持商业的立场,包括其他的三个国会委员会(两个共和党群体和一个众议院民主党派)和两个主要的商业组织(商业圆桌会议和美国商务部)。根据表4.5的图像矩阵,每1个块包含着内部政策交流的高密度,但是在12个块间密度中只有3个块存在经常的联结:支持劳工派的块1发送政策信息给块4的国会部门,而国会部门所在的块4和支持商业的块3也互相交流信息。虽然块模型和MDS分析反映了它们在理念上的清晰区分,但一些组织(例如国会众议院民主党派人士和卡车司机工会)出现在未曾料想的位置上,意味着政策交流在朋友之间和敌对方之间都会发生,这也许可以缓和美国国家劳工政策领域政治对立的紧张程度。

表4.4 美国劳工政策领域核心交流网络的二进制矩阵图

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1	AFL	0	1	1	1	1	0	0	1	1	0	0	0	0	0	1	1	1	1	0	0	1	1	1	1
2	ASCM	1	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	0	0
3	TEAM	0	0	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0	1	1	1	1	1	1	0
4	UAW	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1
5	USW	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0

① 劳联—产联,美国最大的工会联盟。——译者注

续表

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
6	ABC	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	1	0	1	0	1	1	1	1	0
7	AGC	0	0	1	0	0	1	0	1	1	0	1	0	0	0	0	1	0	1	0	1	1	1	1	1
8	BRT	0	0	0	0	0	0	1	0	1	1	1	0	0	0	1	1	1	1	0	1	0	0	1	0
9	CHAM	0	0	1	0	0	1	1	1	0	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0
10	GM	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0
11	NAM	0	0	0	0	1	1	0	1	1	1	0	0	0	0	0	1	0	0	0	1	1	0	1	0
12	NEA	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
13	AARP	1	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1
14	ACLU	1	0	0	0	0	0	0	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0	1
15	NGA	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	1	1	0	1	0	0	0	0
16	DOLS	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0
17	ETA	1	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
18	OSHA	0	0	1	1	0	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
19	NLRB	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
20	WHO	1	0	1	0	0	1	1	1	1	1	0	0	0	0	0	0	0	1	0	0	0	1	1	0
21	HR	0	0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	0	0	1	0	0	0	1	0
22	HD	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	SR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
24	SD	1	0	1	1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0

AFL:美国劳工联合会	ASCM:美国州县市工人联合会	TEAM:卡车司机工会
UAW:汽车工人联合会	USW:钢铁工人联合会	ABC:建筑工人和承包商协会
AGC:通用承包商协会	BRT:商业圆桌会议	CHAM:美国商务部
GM:通用汽车公司	NAM:全美制造商协会	NEA:美国教育协会
AARP:美国退休人员协会	ACLU:美国民权同盟	NGA:全美州长协会
DOLS:劳工部	ETA:劳工部教育和培训管理局	OSHA:职业安全和健康管理局
NLRB:全国劳动关系委员会	WHO:白宫办公室	HR:众议院共和党劳工委员会
HD:众议院民主党劳工委员会	SR:参议院共和党劳工委员会	SD:参议院民主党劳工委员会

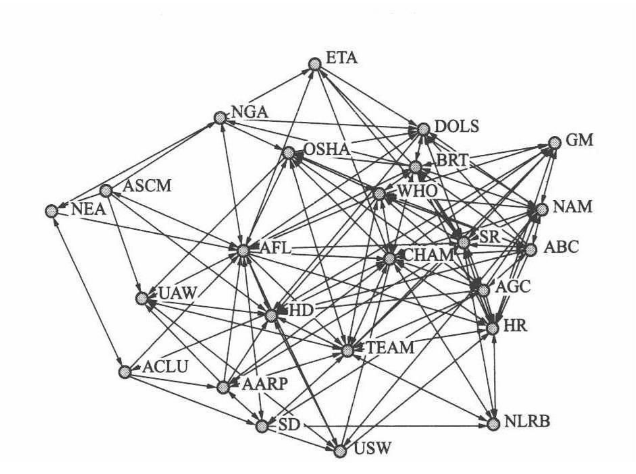


图 4.15　美国国家劳工政策领域关系有向图

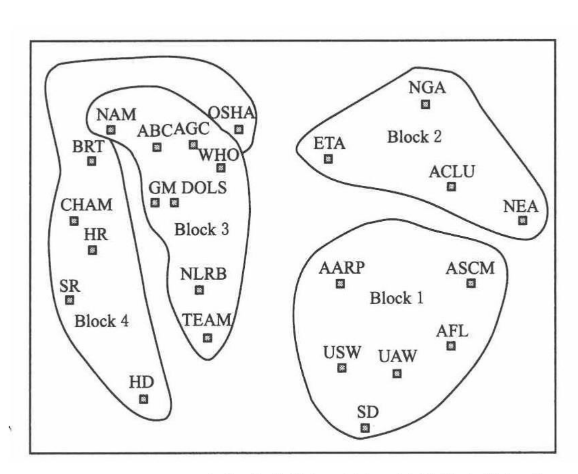


图 4.16　美国国家劳动政策领域关系的块状多维标度

表 4.5 美国劳工政策领域中的核心交流网络的分块阵、密度阵和图像阵

Blocked Matrix

	1	2	24	4	5	13	15	14	12	17	10	3	7	20	6	16	11	19	9	8	21	22	23	18
1 AFL		1	1	1	1		1			1		1				1			1	1	1	1	1	1
2 ASCM	1			1			1		1													1		
24 SD	1			1	1	1						1						1						
4 UAW	1		1		1							1										1		
5 USW	1			1																1	1	1		
13 AARP	1		1									1							1			1	1	
15 NGA	1								1	1				1		1								1
14 ACLU	1		1			1			1										1					
12 NEA	1							1																
17 ETA	1																1		1				1	
10 GM																	1		1	1	1	1	1	
3 TEAM				1		1							1	1				1	1		1	1	1	1
7 AGC			1									1		1	1	1	1		1	1	1	1	1	1
20 WHO	1										1	1	1		1				1	1		1	1	1
6 ABC													1	1		1	1		1	1	1	1	1	1
16 DOLS																			1	1	1	1	1	
11 NAM					1						1			1	1	1			1	1	1		1	
19 NLRB												1							1		1		1	
9 CHAM						1		1			1	1	1	1	1	1	1	1		1	1	1	1	1
8 BRT							1			1	1		1	1		1	1		1				1	1
21 HR											1	1				1	1	1	1	1			1	
22 HD																								
23 SR																								
18 OSHA				1								1	1		1	1								

Density Matrix

	1	2	3	4
Block 1	0.567	0.167	0.125	0.389
Block 2	0.250	0.333	0.094	0.167
Block 3	0.104	0.000	0.393	0.813
Block 4	0.056	0.125	0.458	0.367

Image Matrix

	1	2	3	4
Block 1	1	0	0	1
Block 2	0	1	0	0
Block 3	0	0	1	1
Block 4	0	0	1	1

第5章

网络分析高级方法

在这一章中,我们将讨论有关网络分析的高级方法。我们尤其要讨论网络位置测量、logit 模型 p* 、隶属网络和点阵分析。

第1节 | 网络位置测量

在社会网络分析中，角色和位置是核心概念。利用代数和矩阵方法对互锁角色的分析，始于S. F. 纳德尔(S. F. Nadel, 1957)关于社会结构的理论。在其理论中，他将社会结构视为位置，而该位置拥有与其他位置相关的独特的权利和责任。在第4章讨论过的结构等价，为确认社会网络角色和位置提供了一个基本的方法。然而，它要求结构等价的行为人与同样的其他行为人具有相同的连接模式，这一要求对于该方法的实际运用来说，太过苛刻。虽然放松关系相似性这一要求可以扩展该方法的运用，但是在行为人基于自身的关系模式联合占据位置方面，结构等价依然有着非常严格的定义。作为一个替代性选择，网络研究者创造出多个适用条件没有如此严格的方法来确认网络角色和位置(Borgatti & Everett, 1992; Everett, 1985; Everett, Boyd & Borgatti, 1990; Faust, 1988; Pattison, 1988)。在适用条件严格的降序排列中，结构等价是最严格的，其次是自形和同构等价，再次是正规等价。本部分对后两种方法进行非技术性的一般回顾。虽然自形和同构等价经过一些调整能够被用于有向和赋值图(Wasserman & Faust, 1994:461—502)，但为了便于阐述，我们还是在举例中利用了一个单个关系的无向二进

制图。

同构和自形等价是两个关联性很强的概念，以至于一些研究者认为它们是可以互相替代的(Borgatti & Everett, 1992)。但是，同构等价运用于两个图，而自形等价只能描述一个图中社会关系行为人的关系特征。如果一个图到另一个图的节点有着一对一对应关系，维持了所有节点的邻接关系(例如，同样的点入度和点出度)，那么这两个图就展示了结构上的同质性。也就是说，如果在第一个图中两个节点相连，那么在第二个图中相应的两个点也必然以同样的方式相连(Borgatti & Everett, 1992:11)。每一个图自身是同构的，这就被称为自形形式，也就是一对一的对应关系映射到它自身。当且仅当两个节点与相应的其他位置(但不是同一个节点)连接时，这两个节点就是自形等价(即联合占据着同样的位置)。自形等价节点有相同的图论特征，比如中心度、自我密度、小团体规模(Borgatti & Everett, 1992)。

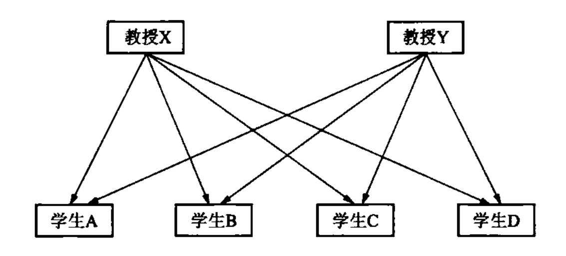
自形等价放宽了结构等价的要求，即在同一位置上的行为人与同样的其他行为人有着相同或非常相似的联结。相反，如果行为人与扮演该位置相同角色的不同行为人有着同样的联结，那么自形等价认同行为人共同占据着一个位置。举一个熟悉的例子：如果两个教授占据着结构等价的位置，那么他们两个必须教同样的一群学生，这实际上是不太可能的。但是，如果说要构成一个自形等价位置，这两个教授所教学生可以不同，但只需要所教学生数量相同即可。学生占据着第二个位置，由教授所教的人来定义。图 5.1 比较了这两种类型的等价，其中从教授到学生的有向连线表示“教学”关系。尽管这两个图显示有两个位置，但自形等价图更好地

抓住了一个理念,即社会角色包含关系的一般模式。现在可以引述另外一个著名的例子:一夫一妻制婚姻。我们期待妻子的位置可由一群女性共同占据,她们与同一群男性有着非多妻制婚姻,但是她们每一个人都单独地与不同的丈夫形成配偶关系。

结构等价性行为人也是自形等价型的,但反之则不一定。如果行为人标签从图中移走,那么自形等价型的节点则不能区分。因此,如果图5.1B中的名字被点所替代,那么这两个子图就无法区分了。博加蒂和埃弗里特(Borgatti & Everett, 1992:16)如此总结这个区别:

> "简要一点来说,我们可以认为,在结构等价方法中,网络或标示图可以代表一个群体的内在结构;因此,在该机构中的行为人的定位就代表了他/她在群体中的位置。相反,在自形等价方法中,利益群体的结构并不是标示图本身;标示图被视为观测结构或表面结构,但表面结构的结构,是潜藏于标示图之下的未标示图。因此在这个'深度结构'中的行为人定位,才真正代表着他/她在群体中的位置。"

A. 结构等价



B. 自形等价

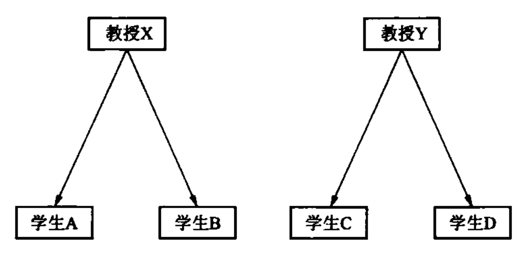


图 5.1 结构等价与自形等价

通过放宽结构等价标准,自形等价法大大便利了对许多社会理论进行实证研究。博加蒂和埃弗里特(1992)认为,有几项采用结构等价操作理论的研究,如果采用自形等价进行分析会更好。例如,他们用此方法回应了伯特(1979)的研究计划。这个研究计划将一个经济体的行业和部门定义为生产相似类型产品、在跨组织网络内占据单一位置的单位集合。博加蒂和埃弗里特(1992:21)则宣称,结构等价性单位,如果按照定义,必须要从同一供应商购买原料,产品必须要销售给同一代理商,实际上这不太可能组成任何符合该定义的部门。但是自形等价单位,则只需要从相似的供应商购买原料,将产品销售给相似的顾客,就能够很好地构成满足要求的行业和部门了。

正规等价是三个常用等价形式中条件最为宽松的一个。它既不要求结构等价中与相同行为人的联结,也不要求自形等价中不可区分的位置。如果行为人与其他也是正规等价的行为人具有同种类型的关系,那么这些行为人就是正规等价的。将这一想法进行概念化的另一条途径是,如果占据一个位置的第一个行为人与占据第二个位置的行为人进行联

结，那么如果第二个行为人构成正规等价，则它必须与第二个位置上的其他行为人有着一个相同的联结（White & Reitz，1983：214）。

所有拥有孩子的母亲都是正规等价，而不论她们孩子数量的多少；反过来也一样。在医院里，医生与他们的患者和护士的关系，也是正规等价的，即使每个医生所属的病人和护士数量不等。正规等价的这种特性使它可能成为社会学家分析社会角色和位置时最为重要的工具。

自形等价和正规等价这两种方法都要求某对行为人与其他在同一种关系具有结构等价性的行为人联系。但是，自形等价和正规等价之间的区别有时也是模糊不清的。自形等价要求未标示图之间必须具有严格的可替代性，但是正规等价并没有要求子图之间具有完全的可替代性。

为了解这种不同，图5.2描述了一个分等级的组织机构图。这个图由四个垂直的层次构成，其中的联结是监督关系。CEO监督三个执行经理（A、B、C），而这三个执行经理又监督下面四个中层管理者（D、E、F、G），再下层就是一些基层的雇员（从H至N）。如果我们忽略雇员层，那么执行经理人B和C就是结构等价，因为他们两个都监督着同样的中层管理者（F和G）。但是，A与B、C不是结构等价的，因为A所监督的管理者不同。然而，这三个执行经理是正规等价型的，因为每一个人都监督着相同数量的中层管理者（两名）。如果我们考虑所有的等级层次，那么B和C也是自形等价型的，因为如果标示被移除，他们的子图都是可以互相替代的。但是，A的子图不能被B和C的子图所替代，因为A的两个中层管理者只监督三个基层雇员，而B和C的子图

每个都有四个雇员。

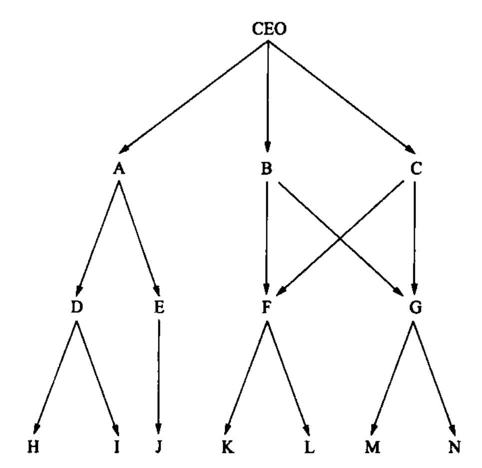


图 5.2 组织等级图的结构等价、自形等价和正规等价

如果仅仅只看两个最低层的关系，那么四个中层管理者没有一个是结构等价的，因为他们都监督着不同的基层雇员。然而，其中三个中层管理者(D、F 和 G)是自形等价的，因为他们各自下属两个雇员，一旦移除标签，是可以完全互相替代的。而 E 不同，它只管理一个雇员。但是，如果按照监督至少一个雇员的标准，所有的中层管理者都符合正规等价标准。图 5.2 清楚地显示，结构等价是最严格的形式，而正规等价是最宽松的形式，自形等价则处于中间。对于确认一般化的网络社会角色，正规等价似乎是一个非常灵活的方法，它宽泛地定义一个社会系统中与其他位置具有相似结构性关系的行为人阶层或类别。

第2节 | logit 模型(p*)

大多数网络分析方法都是描述性的,试图通过数据简化技术去展现一些潜在的社会关系结构,或者通过代数运算来归纳出网络关系的特征。而 p* 模型,这个重要的统计方法,它从对问题的描述提升到探索因果关系,即试图将对偶关系的存在作为个体和图论层次的解释因子的函数。沃瑟曼和帕蒂森(Wasserman & Pattison, 1996)最早提出 p* 模型。这个模型源于 Markov 随机图的早期版本(Frank & Strauss, 1986)——有向图的 log-linear 模型(p1)(Holland & Leinhardt, 1981),而且还运用了伪似然估计的运算法则(Strauss & Ikeda, 1990)。在随后的发展中,p* 模型延伸至分析多变量关系(Pattison & Wasserman, 1999)和赋值关系(Robins, Pattison & Wasserman, 1999)。p* 模型的上述两个运用描述了数据结构及其结果的解释(Anderson, Wasserman & Crouch, 1999; Crouch & Wasserman, 1998),并且提供了一个容易获取的介绍(Wasserman & Robins, 2005)。这一部分将简要地讨论基本 p* 模型,并示范运用于一个拟制的网络(Crouch & Wasserman, 1998:87—101)。虽然这个方法可以运用于多变量关系和赋值图的分析,但这里为简化的缘故,我们只处理一个单一关系的二进制有向图。有兴趣的读

者如果需要进一步学习,可以查阅上述有关文献。

p* 模型与 logistic 回归密切相关。logistic 回归分析二分因变量(作为二进制变量进行编码:$Y=1$ 或 $Y=0$),并假定这些变量是二项式分布。如果最小二乘回归法用于建立作为自变量的线性函数的 Y 的概率模型,那么会产生两个主要的问题:(1)一些预测概率可能会高于 1.0 或低于 0.0,而这些明显是没有意义的概率值;(2)违反了关于误差项的核心假设,特别是违反了等方差(同方差性)且正态分布的假设,以及误差项与自变量无关的假设。为了避免这些问题,logistic 回归模型将概率转换为 logits,具体如下:

$$\text{logit}(Y)=\ln\left(\frac{\Pr(Y=1)}{\Pr(Y=0)}\right)=\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_kX_k \qquad [5.1]$$

利用适当的转换,β 参数可以被解释为自变量对于 Y 的 logit、odds 或概率的效应。odds 转换是最普遍的,因为它比 logit 更直接,比概率方法更简单(Pampel, 2000)。为了获得自变量对 odds 的效应,在上述公式两边以 e 为底取幂:

$$\begin{aligned}\frac{\Pr(Y=1)}{\Pr(Y=0)}&=\exp(\beta_0+\beta_1X_1+\beta_2X_2+\cdots+\beta_kX_k)\\&=e^{\beta_0}e^{\beta_1y_1}\cdots e^{\beta_ky_k}\end{aligned} \qquad [5.2]$$

这个公式表示,自变量对因变量为 1 相对于因变量为 0 的 odds 的影响是可积的。第 i 个自变量对因变量 odds 的净影响是对回归系数取 e 次幂(e^{β_k})。

p* 模型是应用到社会网络数据的 logistic 回归。对于 g 个行为人之间存在二元有向关系的单一关系网络,$g\times g$ 阶矩阵 X 中的单元(i,j)是个二进制数,

$$x_{i,j}=\begin{cases}1 & \text{如果 } i\rightarrow j\\ 0 & \text{其他}\end{cases}$$

沃瑟曼和帕蒂森(1996)提出三个附加矩阵:(1) $X_{i,j}^{+}$ 表示从 i 到 j 强行存在的关系链;(2) $X_{i,j}^{-}$ 表示从 i 到 j 强行缺失的关系链;(3) $X_{i,j}^{C}$ 表示从 i 到 j 的附加关系链。使用这些矩阵,从 i 到 j 的关系链存在且依赖于附加关系链的概率如下所示:

$$\begin{aligned}\Pr(x_{i,j}=1\mid x_{ij}^{C}) &= \frac{\Pr(x=x_{ij}^{+})}{\Pr(x=x_{ij}^{+})+\Pr(x=x_{ij}^{-})}\\ &= \frac{\exp\{\theta Z(x_{ij}^{+})\}}{\exp\{\theta Z(x_{ij}^{+})\}+\exp\{\theta Z(x_{ij}^{-})\}}\end{aligned} \qquad [5.3]$$

其中 θ 表示估计的效应参数的向量,$Z(x_{ij}^{+})$ 和 $Z(x_{ij}^{-})$ 分别表示不同自变量或解释变量的原始数据组成的向量。

从 i 到 j 的关系链存在相对于关系链缺失的 odds 是:

$$\frac{\Pr(x_{i,j}=1\mid x_{ij}^{C})}{\Pr(x_{i,j}=0\mid x_{ij}^{C})}=\frac{\exp\{\theta Z(x_{ij}^{+})\}}{\exp\{\theta Z(x_{ij}^{-})\}} \qquad [5.4]$$

两边取自然对数,然后把公式转换成:

$$\log\left[\frac{\Pr(x_{i,j}=1\mid x_{ij}^{C})}{\Pr(x_{i,j}=0\mid x_{ij}^{C})}\right]=\theta[Z(x_{ij}^{+})-Z(x_{ij}^{-})] \qquad [5.5]$$

这个称为 p* 的公式包括估计的效应参数 θ 向量,和当强行把变量 $x_{i,j}$ 从 1 变成 0 时原始数据构成的向量。

下面我们说明一下几个关于 p* 模型的问题。首先,分

析的单元是一个全幅网络中的对点;因此,对于 N 个行为人组成的有向图,分析单元总数是 N^2-N; 第二,虽然模型看起来像是估计对点的特征,但这实际上是整个图的模型。而且,p* 模型估计观察现在的图作为不同网络特征的函数的条件概率,将这些不同的网络特征(如选择、相互关系和传递性)作为自变量。第三,这些自变量的值是当从 i 到 j 的给定关系链由存在变成缺失时,整个图的变量的合计值的差。例如,假设在某个特定图中相互关系链的总数是 5,如果两个个体之间的关系链存在,如从行为人 4 到行为人 5,然后把从 4 到 5 的关系链变成缺失,那么这个图中的一条相互关系链也将消失。因此,对于两个行为人 (4↔5) 的自变量“相互关系”的值为 1。p* 模型实际上估计的是,如果关系链(x_{ij})是随机形成的,且每个有向关系链发生的概率相互独立时,观察到现在的图形的概率。

为了说明 p* 模型,我们使用克劳奇和沃瑟曼(Crouch & Wasserman, 1998)中的二进制有向图。图 5.3 显示了三个政府研究机构(圆圈表示)和三个私立 R&D 实验室(正方形表示)之间的“项目支持”关系。例如,政府研究机构 1 和政府研究机构 2 之间有相互支持关系。虽然政府研究机构 3 给私立 R&D 实验室 4 提供项目支持,但这个关系链不是互惠的,私立 R&D 实验室 4 对政府研究机构 3 不提供项目支持。

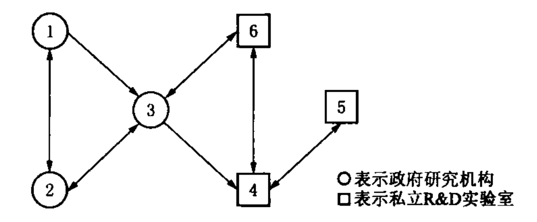


图 5.3 三个政府研究机构和三个私立 R&D 实验室之间“项目支持”的有向图

项目支持似乎更常发生在同一类机构中，也就是说，更常发生在各政府研究机构之间，或各私立实验室之间，而不是机构和私立实验室之间。为了模型化这种“同类”假设和有关关系链存在的其他网络关系的假设，克劳克和沃瑟曼(1998)建模用了五个 p* 参数：(1)整体选择度(θ)；(2)同一类型内的不同选择(θ_w)；(3)相互关系(ρ)；(4)同一类型内的不同相互关系(ρ_w)；(5)传递性(τ_T)，关系是可传递的是指当每次从 i 到 j 存在直接关系链，从 j 到 k 存在直接关系链，那么从 i 到 k 也存在关系链 ($i \rightarrow j$, $j \rightarrow k$, $i \rightarrow k$) (Wasserman & Faust, 1994:164)。在实际的例子中，我们感兴趣的是项目支持是否更常发生在同一类组织之间，互惠的项目支持在同一类组织是否比在不同类组织之间更常发生。因此，要估计的参数向量为：

$$\theta = \{\theta, \theta_w, \rho, \rho_w, \tau_T\}$$

估计有向图中所有两个行为人之间解释变量的向量，包括计算当 i 和 j 之间的关系链从 1 变成 0 时，原始数据的向量 Z(x)的变化。其中：

$Z_\theta = \sum X_{ij}$ 表示解释变量“选择”。

$Z_{\theta_w} = \sum X_{ij}\delta_{ij}$ 表示变量“同一类型内部选择”。

$Z_\rho = \sum_{i<j} x_{ij}x_{ji}$ 表示变量“相互关系”。

$Z_{\rho_w} = \sum_{i<j} x_{ij}x_{ji}\delta_{ij}$ 表示变量“同一类型内部相互关系”。

$Z_{\tau_T} = \sum_{ijk} X_{ij}X_{jk}X_{ik}$ 表示变量“传递性”。

指标变量 δ_{ij} 是个二进制指标，当 i 和 j 是相同类型时，δ_{ij} 为 1，否则为 0。对于有 g 个节点的有向图，p* 模型中有序

数对的总数是 $g(g-1)=g^2-g$。因此,六个组织构成30个有序数对。由于篇幅所限,这里没有呈现所有的 p* 模型参数。有兴趣的读者可以参考安德森(Anderson, 1999)、沃瑟曼和帕蒂森(1996)的相关文献。

表5.1列出了用于估计30个有序数对的 p* 参数的输入数据。下面的例子说明了是如何把这五个公式应用到一个有序对偶关系 (2 → 3) 中来计算输入数据的值的,这个有序对偶关系是表5.1中第7行从行为人2到行为人3的有向关系链。首先,通过当从行为人 i 到行为人 j 的关系链存在时的关系链总数减去当从行为人 i 到行为人 j 的关系链缺失时的关系链总数,从而得到变量"选择":$(\sum X_{ij}^{+}-\sum X_{ij}^{-})$。在这个例子中,对于每对对点,变量"选择"的值为1。因此,变量"选择"实际上是个常数,类似于模型中的截距。第二,同类组织内部的选择(从行为人2到行为人3)为1,因为如果从行为人2到行为人3的关系链存在,则同类组织内的选择总数为9;当这个关系链不存在时,同类组织内的选择总数为8。行为人3和行为人6之间的相互关系链和从行为人3到行为人4的有向关系链没有计算在内是因为这两对对点是不同类型组成的。第三,这个有向图存在五个相互关系链(1↔2, 2↔3, 3↔6, 4↔6, 4↔5)。当从行为人2到行为人3的有向关系链从存在变成缺失时,相互关系链的总数减少到4。因此,从行为人2到行为人3的相互关系变量值为 $5-4=1$。第四,这个有向图存在四条同类组织间的相互关系链(1↔2, 2↔3, 4↔6, 4↔5),不包括行为人3和行为人6之间的相互关系链,因为这对对点是不同类型的。因此,当从行为人2到行为人3的有向关系链从存在变成缺失时,同类组

表5.1 p* 数据

观察个体	i	j	关系链	选择	同类间选择	相互关系	同类间相互关系	传递性
1	1	2	1	1	1	1	1	2
2	1	3	1	1	1	0	0	3
3	1	4	0	1	0	0	0	1
4	1	5	0	1	0	0	0	0
5	1	6	0	1	0	0	0	2
6	2	1	1	1	1	1	1	1
7	2	3	1	1	1	1	1	2
8	2	4	0	1	0	0	0	2
9	2	5	0	1	0	0	0	0
10	2	6	0	1	0	0	0	3
11	3	1	0	1	1	1	1	3
12	3	2	1	1	1	1	1	1
13	3	4	1	1	0	0	0	3
14	3	5	0	1	0	0	0	2
15	3	6	1	1	0	1	0	2
16	4	1	0	1	0	0	0	0
17	4	2	0	1	0	0	0	1
18	4	3	0	1	0	1	0	3
19	4	5	1	1	1	1	1	0
20	4	6	1	1	1	1	1	1
21	5	1	0	1	0	0	0	0
22	5	2	0	1	0	0	0	0
23	5	3	0	1	0	0	0	1
24	5	4	1	1	1	1	1	0
25	5	6	0	1	1	0	0	3
26	6	1	0	1	0	0	0	1
27	6	2	0	1	0	0	0	3
28	6	3	1	1	0	1	0	1
29	6	4	1	1	1	1	1	2
30	6	5	0	1	1	0	0	3

织的相互关系链总数从 4 减少到 3，即减少了 4－3＝1。第五，这个有向图有五种传递的关系 [(1 → 2, 2 → 3, 1 → 3), (1 → 3, 3 → 2, 1 → 2), (2 → 1, 1 → 3, 2 → 3), (3 → 6,

6→4, 3→4), (3→4, 4→6, 3→6)]。当从行为人 2 到行为人 3 的有向关系链从存在变成缺失时，两种传递的关系(1→2, 2→3, 1→3) 和(2→1, 1→3, 2→3) 将消失。因此，这个传递性变量的值为 2。总的来说，对于有序对点(2→3) 的五个计算如下所示：

选择的变化：$\sum X_{ij}^{+} - \sum X_{ij}^{-} = 1$。

同类组织间选择的变化：$\sum X_{ij}^{+}\delta_{ij} - \sum X_{ij}^{-}\delta_{ij} = 9 - 8 = 1$。

相关关系链的变化：$\sum_{i<j}^{+} x_{ij}x_{ji} - \sum_{i<j}^{-} x_{ij}x_{ji} = 5 - 4 = 1$。

同类组织之间相互关系的变化：$\sum_{i<j}^{+} x_{ij}x_{ji}\delta_{ij} - \sum_{i<j}^{-} x_{ij}x_{ji}\delta_{ij} = 4 - 3 = 1$。

传递性的变化：$\sum_{ijk}^{+} X_{ij}X_{jk}X_{ik} - \sum_{ijk}^{-} X_{ij}X_{jk}X_{ik} = 5 - 3 = 2$。

表 5.1 中的数据可以用 SPSS、SAS 和 STATA 等统计软件中的 logistic 回归来分析。因变量是二分变量“关系链”，对自变量——选择，同类组织间的选择，相互关系，同类组织间的相互关系和传递性——进行回归。回归结果如表 5.2 所示。在这个模型中，只有一个解释变量——相互关系链——是显著的。包括从 i 到 j 的有向关系链在内的相互关系链每增加一条，i 发送一条关系链到 j 的 odds 增加将近 41 倍 (exp(3.737) − 1 = 40.97)。控制其他解释变量后，一条相互关系链存在的 odds 是不存在的 odds 的 41 倍多(Anderson et al.，1999:54)。因此，这个有向图中关系链是互惠型的可

能性远远大于不是非互惠型的可能性。然而,互惠和选择对于不同类型的组织是不存在差异的,同类组织内的选择和同类组织内的相互关系都是不显著的。这个结果不支持项目支持在同类组织内比不同组织间更可能发生的假设。

表 5.2 网络关系链的非标准化 logistic 回归

变　量	模型 1	变　量	模型 1
选择	−2.208 (1.197)	传递性	−0.408 (1.197)
同类间选择	2.741 (2.183)	卡方(自由度)	19.488*** (4)
相互关系	3.737* (1.829)	−2 log likelihood	20.893
同类间相互关系	−1.588 (2.921)		

注:括号里的是标准误。* 表示 $p<0.05$,** 表示 $p<0.01$,*** 表示 $p<0.001$(双尾检验)。

在社会网络分析方法中,p* 模型的发展标志着一个非常重要的进步。从此,社会网络分析不仅仅只是描述行为人之间的关系链,还可以解释行为人之间的关系链。模型也能清晰地模拟对图层特征和个体变量的关系链的影响。p* 模型及其类似的模型在网络分析研究人员用来分析实际网络问题时将大有用处。许多实际介绍 p* 模型的教科书将有助于普及这种方法在网络研究人员中的应用。

第3节 | 隶属网络

隶属网络，也称为成员资格网络，表示一组行为人对一系列社会事件的参与(Wasserman & Faust, 1994: 291—343)。人们可以通过对社会活动的联合参与互相联结，或者因他们共同加入某一组织而互相联系。反过来说，社会事件因其共有的参与人而互相联结。隶属网络分析形象地展示了行为人和事件是如何同时相互联系的。

一个隶属网络由两种类型的节点(一组行为人和一系列事件)和每个节点类型之间的一系列关系构成。因此，隶属网络是一种二模网络，这与连接行为人与行为人的一模网络不同。关于隶属网络的实际研究非常多(Wasserman & Faust, 1994: 196)。在第3章，我们讨论了弗里曼等人(1987)关于大学教职工和学生参加九个座谈会的研究。研究者使用隶属网络来表示这些关系，其中参与人是行为人，座谈会是事件。其他一些大家熟悉的例子，有属于志愿者组织的成员、参加抗议活动的社会运动活动者、创立战略联盟的企业，以及签订军事协议的国家等。对二模网络的研究常常具有双重目的：揭示行为人之间通过共同参与事件而形成的关系结构；揭示事件吸引共同的参与人的关系结构。

隶属矩阵和二分图

一个二进制隶属网络可以由一个隶属矩阵来表示，该矩阵记录 g 个行为人在 h 个事件中的参与和缺席。因此，它的维度分别是 g 行和 h 列。如果行为人 i 参与事件 j，那么矩阵中第 i 行第 j 列的单元格的元素等于 1，否则元素为 0。将这个二进制隶属矩阵命名为 A，它的 x_{ij} 值符合下述条件：

$$x_{i,j}=\begin{cases}1 & \text{如果行动者 } i \text{ 与事件 } j \text{ 有关}\\ 0 & \text{其他}\end{cases}$$

矩阵 A 的行总数，也称为行的边际值，记录每一个行为人所参与事件的总数 $\sum_{j=1}^{h} x_{i,j}$。列的边际值（$\sum_{j=1}^{g} x_{i,j}$）表示参与每一个事件的行为人的数量。

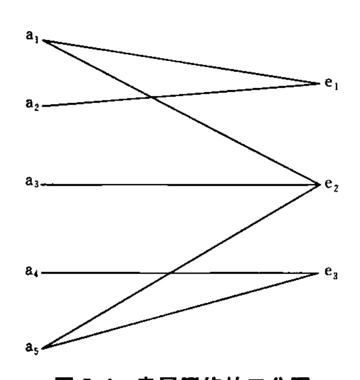


图 5.4 隶属网络的二分图

隶属网络也可以表示为二分图，其中无向线连接排列在图表一端的行为人和排列在图表另一端的事件。重要的是，二分图不允许行为人之间的连线和事件之间的连线。二分

图矩阵(B)在其行和列中包含行为人和事件。假设一个隶属网络有 g 个行为人和 h 个事件,那么 B 的维度为 $(g+h)\times(g+h)$。

图 5.4 是一个隶属网络的二分图,其中有五个行为人和三个事件。线条表示哪一位行为人参与了哪一个事件,同时也表示哪一个事件吸纳了哪一位行为人。虽然在这个图中行为人之间或事件之间没有直接的连线,但是它表明了行为人是如何通过他们共同隶属的具体事件而相互连接的。例如,行为人 1 和行为人 2 都参与了事件 1,而行为人 1、行为人 3、行为人 5 则因事件 2 互相联系在一起。

表 5.3 是对应于图 5.4 的一个二分矩阵,单元格的值为 1 表示一个行为人和一个事件相连。

表 5.3 图 5.4 对应的二分矩阵

	a1	a2	a3	a4	a5	e1	e2	e3	小计
a1	—	0	0	0	0	1	1	0	2
a2	0	—	0	0	0	1	0	0	1
a3	0	0	—	0	0	0	1	0	1
a4	0	0	0	—	0	0	0	1	1
a5	0	0	0	0	—	0	1	1	2
e1	1	1	0	0	0	—	0	0	2
e2	1	0	1	0	1	0	—	0	3
e3	0	0	0	1	1	0	0	—	2
小计	2	1	1	1	2	2	3	2	—

行为人 1 有两个 1 值,表明它参与了两个事件(事件 1 和事件 2)。在右上角象限区内的 5×3 子矩阵,由行的五个行为人和列的三个事件构成,是一个隶属矩阵 A。左下角象限区是 A 的转置,表示为 A′,其维度是 3×5,且 $x_{i,j}=x_{j,i}$。转置的隶属矩阵表明,行中的哪一个事件吸引了列中的哪一

位行为人。二分矩阵的另外两个象限的单元格的值总是为0，因为在二模数据中，行为人之间不直接连接，事件之间也是如此。二分矩阵可以简化为如下形式：

$$X^{A,E} = \begin{bmatrix} 0 & A \\ A' & 0 \end{bmatrix} \qquad [5.6]$$

在表格的边际上（并非矩阵的一部分），行的总数等于列的总数。行为人总数表明它参与事件的数量，而事件总数则表明参与行为人的数量。

将两个处于不同顺序的子矩阵（A 和 A′）相乘，便得到额外的信息，即行为人之间的关系和事件之间的关系，而这在二分矩阵中是没有的。首先，X^A 是一个行为人对点共同出席的赋值的对称矩阵，由两个矩阵相乘获得：

$$X^A = AA' \qquad [5.7]$$

一般来说，隶属网络 A 是一个 $g \times h$ 矩阵，而它的转置 A′是一个 $h \times g$ 矩阵。因此，X^A 总是一个 $g \times g$ 矩阵，它的非对角线单元格的值是行为人 i 和 j 共同参与的事件数量。A 的对角线元素表示每一个行为人参与的事件数量。其次，X^E 是一个事件对点共同参与的赋值的对称矩阵，由两个矩阵相反相乘获得：

$$X^E = A'A \qquad [5.8]$$

这个 $h \times h$ 矩阵的非对角线单元格的值，是同时参与事件 i 和 j 的行为人数量。而 A′矩阵的对角线元素，是参与每一个事件的行为人数量。

案例数据的共同出席矩阵图如下所示：

$$X^A = AA' = \begin{pmatrix} 110 \\ 100 \\ 010 \\ 001 \\ 011 \end{pmatrix} \times \begin{pmatrix} 11000 \\ 10101 \\ 00011 \end{pmatrix} = \begin{pmatrix} 21101 \\ 11000 \\ 10101 \\ 00011 \\ 10112 \end{pmatrix} \quad [5.9]$$

五个行为人和三个事件形成的 X^A 是一个 5×5 的对称矩阵,其非对角线单元格的值表示每一对点参与事件的数量。如第一行所示,行为人 1 与行为人 2、行为人 3 和行为人 5 各参与一个事件,但未与行为人 4 参与事件。对角线表明,行为人 1 和行为人 5 各参与两个事件,而其他人则只参与了一个事件。

案例数据的共同参与矩阵如下所示:

$$X^E = A'A = \begin{pmatrix} 11000 \\ 10101 \\ 00011 \end{pmatrix} \times \begin{pmatrix} 110 \\ 100 \\ 010 \\ 001 \\ 011 \end{pmatrix} = \begin{pmatrix} 210 \\ 131 \\ 012 \end{pmatrix} \quad [5.10]$$

这个 3×3 的对称矩阵的非对角线单元格表示参与每对事件的行为人数量。事件 1 和事件 2 只有一个共同的行为人(其中行为人 1 同时参加了这两个事件),但是事件 1 和事件 3 没有共同的行为人。对角线值反映,事件 2 吸引了三个行为人(行为人 1、行为人 3 和行为人 5)。

在分析的网络层次上,隶属的平均率值是可以从两个乘积矩阵值计算出来的。因为 X^A 的对角线元素是每一个行为人参与事件的数量,因此将它们加起来,然后除以行为人(g)的数量,即得到平均的行为人参与率 $\overline{X}^A = \frac{\sum_{i=1}^{g} x_{i,i}}{g}$。在这个

例子中，平均有 $\left(\frac{7}{5}\right)=1.40$ 个行为人参与每一个事件。同样，将 X^E 的对角线值相加，除以事件(h)总数，即得到平均的事件参与率 $\overline{X}^E=\frac{\sum_{i=1}^{h}x_{i,i}}{h}$。在这个例子中，平均的事件参与率是 $\left(\frac{7}{3}\right)=2.33$ 个行为人。

隶属网络的密度和中心度

密度和中心度是非常重要的基本网络特征，它们也能够适用于隶属网络。正如在第 4 章所讲的，密度测量反映二进制图中联结出现的比例，或者反映赋值图中可观察到的线的平均值。同样地，对隶属网络密度的阐释也依赖于该网络是否为二进制图或赋值图(Wasserman & Faust, 1994:316)。

对于共同参与的 $g\times g$ 对称矩阵 X^A 来说，其非对角线值是每对行为人参与事件的数量，其密度计算公式如下：

$$D^A=\frac{\sum_{i=1}^{g}\sum_{j=1}^{g}X_{ij}^A}{g(g-1)/2}(i<j) \qquad [5.11]$$

在公式 5.11 中，分子表示共同出席矩阵中上三角阵中所有值的总和(即对角线以上的部分，因为下部分三角的值是相同的)。因为我们不可能认为一个行为人与他自己共同参与某一事件，所以我们排除了对角线上的值。分母是无序对点的所有数量，当然也排除了对角线上的值。对于共同参加的 $h\times h$ 对称矩阵 X^E 来说，其非对角线上的值是每一个事

件中参与的行为人数量,其密度计算如下所示:

$$D^E = \frac{\sum_{i=1}^{h}\sum_{j}^{h} X_{i,j}^{E}}{h(h-1)/2}(i<j) \qquad [5.12]$$

在这个例子中,$D^A = 5/10 = 0.50$,意味着每对行为人共同参与了平均半个事件。从事件角度来看,$D^E = 2/3 = 0.67$,表明平均每对事件吸引了 2/3 个行为人。由于乘积矩阵单元格的值可能大于 1,因此密度值也可能大于1.0。因此,每一个密度值不应解释为一个比例,而应该分别解释为每两个行为人所出席的事件平均数(D^A),或每两个事件中的平均参与人数(D^E)。

社会网络分析在行为人的分析层面研究中心度,在图的分析层面研究中心度,这种研究已经进行了数十年(Freeman, 1979; Wasserman & Faust, 1994:第 3 章)。我们在第 4 章讲过,行为人中心度测量一个网络之内行为人的重要性或显性。根据研究者定义重要性和显性的不同,可以将中心度分为四种主要的类型:程度、亲密性、中间性和特征向量中心度。程度中心度反映行为人在网络中活动的积极程度;亲密性中心度测量一个行为人通过网络内最短路径与其他行为人联结的程度;中间性中心度捕捉一个行为人对网络其他行为人之间的信息或资源流的居间程度;特征向量中心度反映行为人与网络中的其他中心行为人的联系程度。浮士德(1997)讨论了这四种中心度测量方法在隶属网络中的运用,但在此由于文本篇幅的限制,我们只讨论隶属网络中程度中心度的使用。

一般认为,程度中心度包含有向联结的总数,因此隶属

网络的行为人程度中心度便是行为人联系的总数，即第 i 个行为人出席所有事件时所接触到的联结数量，由共同出席的矩阵 X^A 的第 i 个行值相加所得：

$$C_D^A(a_i) = \sum_{j=1}^{g} x_{ij}^A (i \neq j) \qquad [5.13]$$

在这个例子中，行为人 1 分别与行为人 2、行为人 3 和行为人 5 一起参与了三个不同事件，因此它的程度中心度是 3。行为人 2 的程度中心度更低，因为它只与行为人 1 参与过第 1 个事件。

同样地，隶属网络中一个事件的程度中心度是事件联系的总数，即第 j 个事件通过所有行为人的参与而得到的联结数量，由共同参与矩阵 X^E 的第 j 个列值相加所得：

$$C_D^E(e_i) = \sum_{j=1}^{h} x_{ij}^E (i \neq j) \qquad [5.14]$$

在这个例子中，三个事件的程度中心度分别为 1、2、1。事件 2 的程度中心度值最高，因为通过至少一个共同的行为人的参与，从而与事件 1 和事件 3 都有了联结。而事件 1 和事件 3 则只与事件 2 共享行为人。

第4节 | 点阵

上面所讨论的隶属网络，一个明显优势就在于可以看见二模数据网络结构的三种类型：行为人—事件关系及其转置，通过共同参与事件而构成的行为人之间的关系，通过吸引共同行为人而构成的事件之间的关系。然而，隶属网络的二分图形式不能揭示行为人和事件节点二者是如何相关的。伽罗瓦点阵(Galois lattices)提供了一个二模数据的一般表述，使得同时查看两个模式之间的关系成为可能，为观察网络的结构特征提供了有价值的研究视角(Freeman & White, 1993; Wasserman & Faust, 1994:326—342)。在实证研究中，该方法运用于下列关系的研究：中国革命年代的村民和政治冲突(Schweizer, 1996)、济贫机构为贫民提供的救济和待遇(Mohr & Duquenne, 1997)，以及1992年巴西弹劾运动中的组织和事件(Mische & Pattison, 2000)。

伽罗瓦点阵

伽罗瓦点阵，因19世纪法国数学家埃瓦里斯特·伽罗瓦(Évariste Galois)的论著而得名，也称为伽罗瓦连接(Galois connection)，是两个有序集合之间的连接。在网络运用

中，伽罗瓦点阵有两个非空集合：行为人集合 A 和事件集合 E。这两个集合通过二模隶属网络连接，表示行为人是如何与事件相连的（反之亦然）。因此，一个伽罗瓦点阵由一个三元组（A, E, I）构成，其中 I 是矩阵 A×E 中的二进制关系。表5.3中二分矩阵的子矩阵 A 是一个行为人—事件矩阵的实例。现在考虑 $P(A)=\{A_1, A_2, \cdots\}$，这是一系列矩阵 A 的集合；$P(E)=\{E_1, E_2, \cdots\}$，这是一系列矩阵 E 的集合。I 关系定义为从 P(A) 至 P(E) 的映射：$B\rightarrow B\uparrow$：

$$B\uparrow=\{e\in E\mid(a, e)\in I,\text{所有 } a\in A\} \quad [5.15]$$

这个数学公式表明，映射图能够识别某一特定行为人或某些行为人所隶属的所有事件。尤其是，↑映射从一系列行为人到一系列事件，其中这些行为人隶属于这些事件（Wasserman & Faust，1994：330）。例如，表 5.3 中的子矩阵 A 表示，行为人 1 隶属于事件 1 和事件 2，而行为人 1 和行为人 2 都隶属于事件 1。

反过来，这个映射能够从 P(E) 至 P(A)：$F\rightarrow F\downarrow$

$$F\downarrow=\{a\in A\mid(a, e)\in I,\text{所有 } e\in E\} \quad [5.16]$$

公式 5.16 意味着映射应当确认一个具体事件所吸引的所有行为人。表 5.3 中的子矩阵 A 显示，事件 2 吸引了行为人 1、行为人 3 和行为人 5，但事件 2 和事件 3 只吸引了行为人 5。

将两个映射图合并起来，伽罗瓦点阵展现了一系列行为人是如何隶属于一系列事件的。在图表中的每一个点，既表示一系列行为人，也表示一系列事件。为了解释伽罗瓦点阵，需要在一个一致的方向上移动：(1)为了跟踪出席事件的行为人，只需要向上移动，将点阵从底端升至顶端；(2)为了

跟踪吸引了行为人的事件,只需要向下移动,将点阵下降。依照惯例,点阵上面部分涵盖了行为人集合 A 的全部要素,下面部分涵盖了行为人集合 E 的全部要素。图 5.5 展示了表 5.3 当中隶属矩阵的伽罗瓦点阵图。这个图将三个事件标示为 A、B、C,将五个行为人标示为 1、2、3、4、5。该点阵底部聚集了最多的事件。沿着点阵向上移动,会逐渐遇到行为人集合的增大和事件集合的缩小。在底端的起始点,包含所有的事件(ABC),但行为人集是空的(ϕ),因为没有人参与所有的事件。向上一点,点阵图显示,事件 A 和事件 B 共享行为人 1,事件 B 和事件 C 共享行为人 5。再往上,事件集合变得更小,但行为人集合变大。事件 A 吸引了行为人 1 和行为人 2,而事件 B 吸引了行为人 1、行为人 3 和行为人 5,事件 C 吸引了行为人 4 和行为人 5。点阵最顶端罗列了所有的行为人但没有任何事件,这表明没有事件吸引所有的网络行为人。一般来说,如果接收到从事件处下降之线条上行为人是入射的,那么该行为人隶属于该事件。在以事件 A 和事件 B 为出发点的线条上,行为人 1 是入射的,表明行为人 1 隶属于这两个事件。相反,在从行为人处升起之线条上是入射的事件,吸引了这些行为人。例如,事件 C 在由行为人 5 处升起之线条上是入射的,表明事件 C 吸引了行为人 5。事件 B 收到来自行为人 1 和

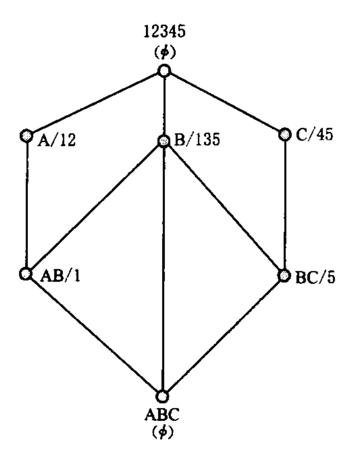


图 5.5 隶属网络的伽罗瓦点阵

行为人5的线条，表明这两个行为人都参与了事件B。

伽罗瓦点阵也表示一些事件之间的隶属模型和行为人之间的隶属模型。例如，图5.5显示，行为人2没有参与任何无行为人1参加的其他事件(行为人2只参加了事件A，行为人1也隶属其中)。同样，行为人4只参加了事件C，该事件也吸引了行为人5。相反，行为人1或行为人5都没有限制自己参加同样吸引了行为人2或行为人4的事件。也就是说，对于行为人1或行为人5来讲，参与某个事件并不以其他行为人的参与为条件。从事件的角度来看，图5.5也显示，三个事件集合中的行为人有重叠但不相同，在此意义上三个事件都包含不同的行为人集合。含有更多行为人和事件的更复杂网络，也许会发现一些包容性结构，其中在某一个事件中出现的行为人也出现在其他的事件里(Freeman & White, 1993)。因此，所有三种连接类型——行为人—事件关系、行为人—行为人关系、事件—事件关系——的数据都能够用伽罗瓦点阵进行精确重组。

使用伽罗瓦点阵对三种关系类型的结构特点进行同步分析，尽管这是一个非常明显的优势，但是它的应用只限于表述小型的二模网络数据集合。关于这方面，伽罗瓦点阵类似于图，其主要的运用是展示数据，而不是还原数据。大型数据集合，一般包含高度复杂的结构，会迅速搅乱伽罗瓦点阵的可视性。即使利用简化符号，如弗里曼和怀特(1993)所建议的那样，观察者依然很难看清楚复杂的伽罗瓦点阵的图像。

网络小团体的伽罗瓦点阵

网络分析者认为，一些统计的或代数的数据还原方法可

以被用于简化伽罗瓦点阵的可视化图像(Duquenne, 1996; White, 1996)。弗里曼(1996)认为,伽罗瓦点阵能够帮助展示社会行为人之间的小团体。经典的卢斯—佩里(Luce-Perry)小团体定义假定存在一个关于社会关系 R 的二进制对称平方矩阵(A^2)。小团体 C 是一个含有三个或多个行为人的最大子集,其中每对行为人都由 R 连接。"最大"意味着没有小团体能涵盖在一个更大的小团体中。然而,在实践中,小团体往往太小、太多或者太过于重叠,以至于不能反映社会群体的直观结构(Freeman, 1996:174)。伽罗瓦点阵方法运用于小团体结构的可视化处理,需要用小团体集替代事件集。之前我们讲到过,伽罗瓦点阵由一个三元组(A、E 和 I)构成,其中 A 是行为人集,E 是事件集,I 是 A×E 中的二进制关系。如果将伽罗瓦点阵应用于小团体,那么伽罗瓦点阵将被重新定义为另一个三元组(A、C 和 I),其中 A 和 I 的意义不变,C 表示小团体的集合。

遵循事件的伽罗瓦点阵的相似布局,小团体伽罗瓦点阵将较大的小团体集合放置在下部,将较大的行为人集合放置在上部。假定图 5.5 现在表示的是一个小团体伽罗瓦点阵,而不是一个事件伽罗瓦点阵,那么在底部有三个小团体 ABC,其零值表示没有行为人属于这三个小团体。沿着点阵往上走,我们可以发现,小团体越少,行为人集合越大。行为人 1 同时属于小团体 A 和小团体 B,而行为人 5 同时属于小团体 B 和小团体 C。继续往上走,可以看到小团体 A 包含行为人 1 和行为人 2,小团体 B 包含行为人 1、行为人 3 和行为人 5,而小团体 C 包含行为人 4 和行为人 5。行为人的全体集合位于图的顶端,是个空集,表示没有一个小团体同时包

含所有五个行为人。

弗里曼(1996)讨论过小团体网络伽罗瓦点阵的几个重要结构特征。首先，两个重叠小团体由下降线连接，交会于点阵图低处某个标签点，而两个非重叠小团体则仅在无标签泛下限处与空集(ϕ)连接。假设图 5.5 反映的是一个小团体网络的伽罗瓦点阵，小团体 A 和小团体 B 交会于一个标有行为人 1 的较低点，表示两个小团体出现重叠，共享了行为人 1。相反，小团体 A 和小团体 C 没有重叠，表示它们的交会点是在无标签的泛下限点。

其次，弗里曼(1996)利用多种重要维度(如链、长度、高度和深度)，对小团体点阵的个人行为人的位置按特点进行了归纳。链是指由一个元素引至另一个元素的全部上升线或下降线所组成的序列。链的长度是链所包含的线条的数量。行为人的高度是从泛下限点上升至该行为人处的链的长度。相反，行为人的深度是从泛上限点下降至该行为人处的链的长度。因此，处于点阵低处的行为人具有较大的深度和较小的高度，而处于点阵顶部的行为人具有较大的高度但是较小的深度。深度较大但高度较小的行为人深嵌在网络中。他们参与多个小团体，但他们与这些小团体的隶属关系独立于其他行为人的隶属关系。从某种意义上来说，他们构成了这个小团体的核心成员。相反，高度较大但深度较小的行为人则往往浮在网络表层。他们与某个小团体的隶属关系往往取决于其他行为人的隶属情况。因此，他们属于这个小团体的边缘人物。将这些观点套用于图 5.5，我们就能够确定行为人 1 和行为人 5 分别是小团体 AB 和小团体 BC 的核心成员，而行为人 2、行为人 3 和行为人 4 则是这些小团体的边缘人物。

对应分析

上面讲的点阵分析是运用代数方法展示隶属网络的代表，而对应分析则利用度量技术来联合展现隶属网络中所有的行为人和事件(Faust, 2005; Wasserman & Faust, 1994: 291—343)。对应分析主要通过奇异值分解(Singular Value Decomposition, SVD)的数学方法进行。我们这里只简要地介绍一下 SVD,集中谈论与隶属网络矩阵有直接关系的几个事项(欲知详情,请参阅 Strang, 1988)。

SVD 是矩阵 A 的分解,其维度有 $g \times h$:

$$A = X \Lambda Y^{T} \qquad [5.17]$$

这个公式含有 Λ,它表示奇异值的对角线矩阵 $\{\lambda_K\}$;X 表示大小为 $g \times h$ 的左奇异向量矩阵;Y 表示大小为 $h \times h$ 的右奇异向量矩阵。因此,左奇异向量指矩阵 $g \times h$,右奇异向量指矩阵 $h \times h$。奇异值的数量也被称为秩,通常表示为 W。SVD 利用秩 W 来度量视图中的行为人和事件,将它们排列在 A 的合适位置中。

对应分析中的 SVD 包括对标准化后的矩阵 A 的分解。标准化有两种方法:(1)将原始矩阵 A 的元素除以行和列边际总数乘积的平方根;(2)计算矩阵 A 和两个对角线矩阵 $R^{-\frac{1}{2}}$ 和 $C^{-\frac{1}{2}}$ 的乘积(Faust, 2005),其中

$$R^{-\frac{1}{2}} = \mathrm{diag}\left(\frac{1}{\sqrt{a_{i+}}}\right) \qquad [5.18]$$

$$C^{-\frac{1}{2}} = \mathrm{diag}\left(\frac{1}{\sqrt{a_{+j}}}\right) \qquad [5.19]$$

将三个矩阵相乘，得到矩阵 A 的标准化形式，标准化矩阵 A 的获得也可以通过如下方法：用矩阵 A 的元素除以行和边际总数乘积的平方根。对应分析包含乘积矩阵的 SVD：

$$R^{-\frac{1}{2}}AC^{-\frac{1}{2}} = X \wedge Y^T \quad [5.20]$$

对应分析获得三个信息集合：行矩阵的 g 值集合 $U = \{u_{ik}\}$，其中 $i = 1, 2, \cdots g$，$k = 1, 2, \cdots w$；列矩阵的 h 值集合 $V = \{v_{jk}\}$，其中 $j = 1, 2, \cdots h$，$k = 1, 2, \cdots w$；奇异值 $\wedge = \{\lambda_K\}$，其中 $k = 1, 2, \cdots w$，表示每一个维度的重要性。为了获得一个所有行和列元素的联合展示，对应分析将行为人的值计算为其所隶属的事件加权平均值的函数，将事件的值计算为它的行为人值的加权平均值的函数：

$$\lambda_k u_{ik} = \sum_{j=1}^{h} \frac{a_{ij}}{a_{i+}} v_{jk} \quad [5.21]$$

$$\lambda_k v_{jk} = \sum_{j=1}^{g} \frac{a_{ij}}{a_{+j}} u_{ik} \quad [5.22]$$

在上述两个公式里，a_{ij} 是原始矩阵 A 中第 i 行和第 j 列的单元格的值。

为了讲解的便利，我们利用前面的隶属网络数据：五个行为人和三个事件。表 5.4 表示原始版和标准化版的矩阵 A。表 5.5 列出了标准化矩阵 A 的 SVD，它得到三个数据集合 u_{ik}、v_{jk} 和 λ_K。行为人值是其所隶属的事件值的一个函数。同样，事件值是其所吸引的行为人值的一个函数。具体来讲，在第一维度行为人 1 的值（−0.661）是由该行为人隶属的两个事件的加权平均值除以奇异值（λ_1）得出：$\dfrac{\frac{1}{2} \times (-1.146) + \frac{1}{2} \times 0}{0.866} = -0.661$。同样，在维度 1 中事件

1 的值(−1.146)是由事件所吸引之行为人(1 和 2)的加权平均值得来的：$\dfrac{\frac{1}{2}\times(-0.661)+\frac{1}{2}\times(-1.323)}{0.866}=-1.146$。那么，你能计算出其他行为人和事件的值吗？

表 5.4 隶属网络的矩阵 A

	e1	e2	e3
隶属矩阵 A			
a1	1	1	0
a2	1	0	0
a3	0	1	0
a4	0	0	1
a5	0	1	1
标准化矩阵 A			
a1	0.5	0.408	0
a2	0.707	0	0
a3	0	0.577	0
a4	0	0	0.707
a5	0	0.408	0.5

图 5.6 给出了隶属网络的对应分析。括号内的数是两个维度的值，等于表 5.5 中的数。X 轴和 Y 轴分别表示第一和第二维度(λ)。在图中，行为人 1 隶属于事件 1 和事件 2，行为人 5 隶属于事件 2 和事件 3，处于两个维度的中心，因此位于图表的中央。事件 2 位于第一维度的中心，主要是因为它吸引了大多数行为人(1、3 和 5)，但是它离第二维度的中心要比事件 1 和事件 3(每个事件只吸引了两个行为人)远一点点。行为人 3 位于第一维度的中心，因为它隶属于中心事件(事件 2)，但是行为人 3 在第二维度中的位置却是最边缘的，也许是由于它只参与了一个事件。行为人 2 和行为人 4

在两个维度中都处于边缘，反映他们只参与了一个事件。同样，事件1和事件3在两个维度中也是位于边缘，表明它们比事件2吸引的行为人要少一些。

表5.5　隶属网络的对应分析主坐标

维度(w)	1	2
	行为人/行值(u_{ik})	
A1	−0.661	−0.144
A2	−0.323	0.866
A3	0.000	−1.155
A4	1.323	0.866
A5	0.661	−0.144
	事件/列值(v_{jk})	
E1	−1.146	0.559
E2	0	−0.745
E3	1.146	0.559
奇异值(λ_k)	0.866	0.645

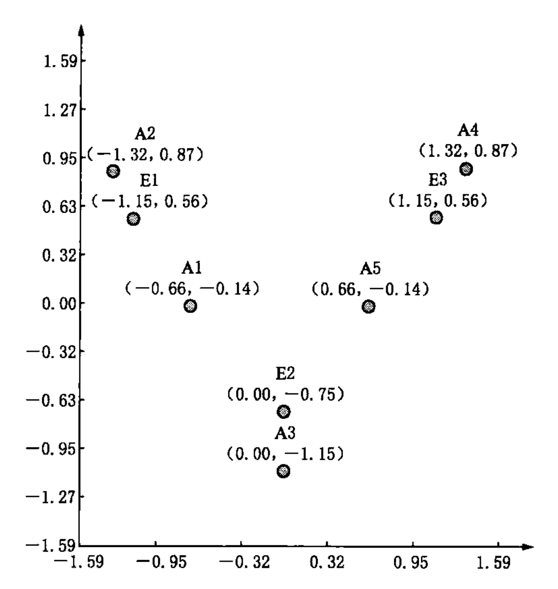


图5.6　5-行为人和3-事件隶属网络的对应分析

结论

在前面有关章节中,我们讨论了一些网络分析的方法(聚类分析、MDS 和相应分析)。这些方法在用于非网络数据的分析时具有明显的优势。例如,韩(2003)利用加权 MDS 法,分析了 1993 年综合社会调查数据中受访者的音乐爱好。虽然收集社会网络数据耗时费力,尤其是采用全幅网络设计时,但社会调查数据却比较容易获得,可以用这些网络分析方法对它进行分析,以揭示潜在的关系。社会网络分析为研究者提供了有力的分析概念和工具方法。最近有一项研究非常娴熟地运用了聚类分析方法,将中欧东部地区 1586 个市民社会网站根据 13 个变量分类成六个维度(Vedres, Bruszt & Stark, 2004)。另一项研究(Mohr, 1998)则直接讨论如何将 MDS、聚类和对应分析等方法用于对复杂结构进行简化,对不同文化背景之间的一些深层结构逻辑进行可视化处理。这些研究展现了多种社会网络方法的熟练应用,也正好表达了我们本书所追求的最终目的——鼓励读者朋友努力学习社会网络分析,以洞悉结构关系之谜。

参考文献

Adamic, L. A. , & Adar, E. (2003). Friends and neighbors on the Web. *Social Networks*, *25*, 211—230.

Alba, R. D. (1973). A graph-theoretic definition of a sociometric clique. *Journal of Mathematical Sociology*, *3*, 113—126.

Aldenderfer, M. S. , & Blashfield, R. K. (1984). *Cluster analysis*. Beverly Hills, CA: Sage.

Alderson, A. , & Beckfield, J. (2004). Power and position in the world city system. *American Journal of Sociology*, *109*, 811—851.

Alexander, M. C. , & Danowski, J. (1990). Analysis of an ancient network: Personal communication and the study of social structure in a past society. *Social Networks*, *12*, 313—335.

Allison, P. (2001). *Missing data*. Thousand Oaks, CA: Sage.

Anderson, C. , Wasserman, S. , & Crouch, B. (1999). A p* primer: Logit models for social networks. *Social Networks*, *21*, 37—66.

Anderson, J. G. , & Jay, S. (1985). Computers and clinical judgement: The role of physician networks. *Social Science and Medicine*, *20*, 969—979.

Bailey, S. , & Marsden, P. (1999). Interpretation and interview context: Examining the General Social Survey name generator using cognitive methods. *Social Networks*, *21*, 287—309.

Baker, W. , & Faulkner, R. (1993). The social organization of conspiracy. *American Sociological Review*, *58*, 837—860.

Barabási, A. -L. (2002). *Linked: The new science of networks*. Cambridge, MA: Perseus.

Barnes, J. (1954). Class and committees in a Norwegian island parish. *Human Relations*, *7*, 39—58.

Barnes, J. A. (1979). Network analysis: Orienting notion, rigorous technique or substantive field of study. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives on social network analysis* (pp. 403—423). New York: Academic Press.

Batchelder, E. (2002). Comparing three simultaneous measurements of a sociocognitive network. *Social Networks*, *24*, 261—277.

Bearman, P. S. , Moody, J. , & Stovel, K. (2004). Chains of affection:

The structure of adolescent romantic and sexual networks. *American Journal of Sociology*, *110*, 44—91.

Beauchamp, M. (1965). An improved index of centrality. *Behavioral Science*, *10*, 161—163.

Benson, M. A. (2004). Dyadic hostility and the ties that bind: State-to-state versus state-to-system security and economic relationships. *Journal of Peace Research*, *41*, 659—676.

Berman, F., & Brady, H. (2005). *Final report: NSF SBE-CISE workshop on cyberinfrastructure and the social sciences*. Washington, DC: National Science Foundation.

Bernard, R. H., Johnsen, E., Killworth, P., McCarty, C., Shelley, G., & Robinson, S. (1990). Comparing four different methods for measuring personal social networks. *Social Networks*, *12*, 179—215.

Bernard, R. H., & Killworth, P. (1977). Informant accuracy in social network data Ⅱ. *Human Communications Research*, 4, 3—18.

Bernard, R. H., Killworth, P., & Sailer, L. (1981). Summary of research on informant accuracy in network data, and on the reverse small world problem. *Connections*, *4*(2):11—25.

Bernard, R. H., Killworth, P., Sailer, L., & Kronenfeld, D. (1984). The problem of informant accuracy: The validity of retrospective data. *Annual Review of Anthropology*, *13*, 495—517.

Blau, P. M., Ruan, D., & Ardelt, M. (1991). Interpersonal choice and networks in China. *Social Forces*, *69*, 1037—1062.

Boorman, S., & White, H. (1976). Social structure from multiple networks. Ⅱ. Role structures. *American Journal of Sociology*, *81*, 1384—1446.

Borgatti, S., & Everett, M. (1992). Notions of position in social network analysis. *Sociological Methodology 22*, 1—35.

Borgatti, S., Everett, M., & Freeman, L. (2004). *UCINET VI software for social network analysis*. Natick, MA: Analytic Technologies.

Borgatti, S., Everett, M., & Shirey, P. (1990). LS sets, lambda sets and other cohesive subsets. *Social Networks*, *12*, 337—357.

Borgatti, S., & Molina, J.-L. (2005). Towards ethical guidelines for network research in organizations. *Social Networks*, *27*, 107—117.

Borgatti, S. P. (2003). The key player problem. In R. Breiger, K. Carley, & P. Pattison(Eds.), *Dynamic social network modeling and analysis*:

Workshop summary and papers (pp. 241—252). Washington, DC: National Academy of Sciences Press.

Brandes, U., Kenis, P., Raab, J., Schneider, V., & Wagner, D. (1999). Explorations into the visualization of policy networks. *Journal of Theoretical Politics*, *11*, 75—106.

Brewer, D. D. (2000). Forgetting in the recall-based elicitation of personal and social networks. *Social Networks*, *22*, 29—43.

Brewer, D. D., & Webster, C. (1999). Forgetting of friends and its effects on measuring friendship networks. *Social Networks*, *21*, 361—373.

Burgin, R. (1995). The retrieval effectiveness of five clustering algorithms as a function of indexing exhaustivity. *Journal of the American Society for Information Science*, *46*(8), 562—572.

Burt, R. (1983). Range. In R. S. Burt & M. Minor(Eds). *Applied network analysis: A methodological introduction*. Beverly Hills, CA: Sage.

Burt, R. (1985). General social survey network items. *Connections*, *8*, 119—123.

Burt R. (1987). A note on missing network data in the General Social Survey. *Social Networks*, *9*, 63—73.

Burt, R. S. (1978). Cohesion versus structural equivalence as a basis for network subgroups. *Sociological Methods and Research*, *7*, 189—212.

Burt, R. S. (1979). Disaggregating the effect on profits in manufacturing industries of having imperfectly competitive consumers and suppliers. *Social Science Research*, *8*(2), 120—143.

Burt, R. S. (1983). Distinguishing relational contents. In R. S. Burt & M. J. Minor (Eds.), *Applied network analysis* (pp. 35—74). Beverly Hills, CA: Sage.

Burt, R. S. (1992). *Structural holes: The social structure of competition*. Cambridge, MA: Harvard University Press.

Campbell K. E., & Lee, B. (1991). Name generators in surveys of personal networks. *Social Networks*, *13*, 203—221.

Campbell, K. E., Marsden, P., & Hurlbert, J. (1986). Social resources and socioeconomic status. *Social Networks*, *8*, 97—117.

Carley, K., & Krackhardt, D. (1996). Cognitive inconsistencies and nonsymmetric friendship. *Social Networks*, *18*, 1—27.

Carrington, P., Scott, J., & Wasserman, S. (Eds.). (2005). *Models and methods in social network analysis*. New York: Cambridge University

Press.

Casciaro, T. (1998). Seeing things clearly: Social structure, personality, and accuracy in social network perception. *Social Networks*, *20*, 331—351.

Cohen, S., Frank, E., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr. (1998). Types of stressors that increase susceptibility to the common cold in adults. *Health Psychology*, *17*, 214—223.

Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American Journal of Sociology*, *91*, 1309—1335.

Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, *28*, 1557—1575.

Cross, R., Liedtka, J., & Weiss, L. (2005). A practical guide to social networks. *Harvard Business Review*, *83*(3), 124—132.

Crouch, B., & Wasserman, S. (1998). A practical guide to fitting social network models via logistic regression. *Connections*, *21*, 87—101.

Davis, J. A. (1979). The Davis/Holland/Leinhardt studies: An overview. In P. W. Holland & S. Leinhardt(Eds.). *Perspectives on social network research*(pp. 51—62). New York: Academic Press.

Degenne, A., & Forse, M. (1990). *Introducing social networks*. Thousand Oaks, CA: Sage.

Diani, M., & McAdam, D. (Eds.). (2003). *Social movements and networks: Relational approaches to collective action*. New York: Oxford University Press.

Doreian, P., Batagelj, V., & Ferligoj, A. (2005). *Generalized blockmodeling*. Cambridge. UK: Cambridge University Press.

Doreian, P., & Woodard, K. (1992). Fixed list versus snowball selection of social networks. *Social Science Research*, *21*, 216—233.

Doreian, P., & Woodard, K. (1994). Defining and locating cores and boundaries of social networks. *Social Networks*, *16*, 267—293.

Dunbar, R. I. M., & Spoor, M. (1995). Social networks, support cliques and kinship. *Human Nature*, *6*, 273—290.

Duquenne, V. (1996). On lattice approximations: Syntactic aspects. *Social Network*, *18*, 189—199.

Emirbayer, M. (1997). Manifesto for a relational sociology. *American Jour-*

nal of Sociology, *103*, 281—317.

Erickson, B. H. (1996). Culture, class, and connections. *American Journal of Sociology*, *102*, 217—251.

Erickson, B. H. (2004). The distribution of gendered social capital in Canada. In H. Flap & B. Volker(Eds.), *Creation and returns of social capital*: *A new research program*(pp. 27—51). New York: Routledge.

Everett, M. (1985). Role similarity and complexity in social networks. *Social Networks*, *7*, 353—359.

Everett, M., Boyd, J., & Borgatti, S. (1990). Ego-centreed and local roles: A graph theoretic approach. *Journal of Mathematical Sociology*, *15*, 163—172.

Fafchamps, M., van der Leij, M. J., & Goyal, S. (2006). *Scientific networks and co-authorship*(*No. 256*). University of Oxford, Department of Economics, Economics Series Working Papers.

Faust, K. (1988). Comparison of methods for positional analysis: Structural and general equivalences. *Social Networks 10*, 313—341.

Faust, K. (1997). Centrality in affiliation networks. *Social Networks 19*, 157—191.

Faust, K. (2005). Using correspondence analysis for joint displays of affiliation networks. In P. J. Carrington, J. Scott, & S. Wasserman(Eds.), *Models and methods in social network analysis* (pp. 117—147). New York: Cambridge University Press.

Faust, K., & Skvoretz, J. (2002). Comparing networks across space and time, size and species. *Sociological Methodology*, *32*, 267—299.

Feld, S. L., & Carter, W. C. (2002). Detecting measurement bias in respondent reports of personal networks. *Social Networks*, *24*, 365—383.

Feldman-Savelsberg, P., Ndonko, F. L., & Yang, S. (2005). How rumor begets rumor: Collective memory, ethnic conflict, and reproductive rumors in Cameroon. In G. A. Fine, V. Campion-Vincent, & C. Heath (Eds), *Rumor mills*: *The social impact of rumor and legend* (pp. 141—159). New York: Transaction Press.

Fischer, C. S. (1982). What do we mean by "friend"? An inductive study. *Social Networks 3*, 287—306.

Fischer, C. S., & Shavit, Y. (1995). National differences in network density: Israel and the United States. *Social Networks 17*, 129—145.

Frank, O. (2003). *KliqFinder for Windows*, *version 0. 05*. East Lansing:

Michigan State University.

Frank, O. (2005). Network sampling and model fitting. In P. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 31—45). New York: Cambridge University Press.

Frank, O., & Snijders, T. A. B. (1994). Estimating the size of hidden populations using snowball sampling. *Journal of Official Statistics*, *10*, 53—67.

Frank, O., & Strauss, D. (1986). Markov graphs. *Journal of the American Statistical Association 81*, 832—842.

Freeman, L. (1977). A set of measures of centrality based upon betweenness. *Sociometry*, *40*, 35—41.

Freeman, L. (1979). Centrality in social networks: I. Conceptual clarification. *Social Networks*, *1*, 215—239.

Freeman, L. C. (1992). Filling in the blanks: A theory of cognitive categories and the structure of social affiliation. *Social Psychology Quarterly*, *55*, 118—127.

Freeman, L. C. (1992). The resurrection of cliques: Application of Galois lattices. *Bulletin de Methodologie Sociologique 37*, 3—24.

Freeman, L. C. (1996). Some antecedents of social network analysis. *Connections*, *19*, 39—42.

Freeman, L. C. (2000). Visualizing social networks, *Journal of Social Structure*, *1*, 1—15.

Freeman, L. C. (2004). *The development of social network analysis: A study in the sociology of science*. Vancouver, Canada: Empirical Press.

Freeman, L. C. (2005). Graphic techniques for exploring social network data. In P. J. Carrington, J. Scott, & S. Wasserman (Eds.), *Models and methods in social network analysis* (pp. 248—270). Cambridge, MA: Cambridge University Press.

Freeman, L. C., Borgatti, S., & White, D. (1991). Centrality in valued graphs: A measure of betweenness based on network flow. *Social Networks*, *13*, 141—154.

Freeman, L. C., Romney, K., & Freeman, S. (1987). Cognitive structure and informant accuracy. *American Anthropologist*, *89* (2), 310—325.

Freeman, L. C., & Thompson, C. R. (1989). Estimating acquaintanceship volume. In M. Kochen (Ed.), *The small world* (pp. 147—158). Norwood, NJ: Ablex.

Freeman, L. C. , & Webster, C. (1994). Interpersonal proximity in social and cognitive space. *Social Cognition*, *12*(3), 223—247.

Freeman, L. C. , & White, D. (1993). Using Galois lattices to represent network data. *Sociological Methodology*, 23, 127—146.

Galaskiewicz, J. (1979). The structure of community organizational networks. *Social Forces*, *57*, 1346—1364.

Goodman, L. (1961). Snowball sampling. *The Annals of Mathematical Statistics*, *32*, 148—170.

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, *78*, 1360—1380.

Han, S. (2003). Unraveling the brow: What and how of choice in musical preference. *Sociological Perspectives*, *46*, 435—459.

Harary, F. (1969). *Graph theory*. Reading, MA: Addison-Wesley.

Hargens, L. L. (2000). Using the literature: Reference networks, reference contexts, and the social structure of scholarship. *American Sociological Review*, *65*, 846—865.

Harrisson, D. , Laplante, N. , & St-Cyr, L. (2001). Cooperation and resistance in work innovation networks. *Human Relations*, *54*, 215—255.

Heckathorn, D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social Problems*, *44*, 174—199.

Heider, F. (1958). *The psychology of interpersonal relations*. New York: Wiley.

Holland, P. W. , & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association*, *76*, 33—65.

Huisman, M. , & van Duijn, M. A. J. (2005). Software for social network analysis. In P. J. Carrington, J. Scott, & S. Wasserman(Eds.), *Models and methods in social network analysis* (pp. 270—316). New York: Cambridge University Press.

Johnson, J. , & Orbach, M. (2002). Perceiving the political landscape: Ego biases in cognitive political networks. *Social Networks*, *24*, 291—310.

Jolly, A. M. , Muth, S. Q. , Wylie, J. L. , & Potterat, J. J. (2001). Sexual networks and sexually transmitted infections: A tale of two cities. *Journal of Urban Health*, *78*(3), 433—445.

Jones, L. M. , & Fischer, C. S. (1978). *Studying egocentric network by mass survey*. Working Paper at the Institute of Urban and Regional De-

velopment, University of California, Berkeley.

Kang, S.-Y., Deren, S., Andia, J., Colon, H. M., & Robles, R. (2005). Egocentric HIV risk networks among Puerto Rican crack users in New York and in Puerto Rico: Impact on sex risk behaviors over time. *AIDS Education and Prevention*, *17*, 53—67.

Kenis, P., & Knoke, D. (2002). How organizational field networks shape interorganizational tie-formation rates. *Academy of Management Review*, *27*, 275—293.

Killworth, P., Johnsen, E., Bernard, R., Shelley, G., & McCarty, C. (1990). Estimating the size of personal networks. *Social Networks*, *12*, 289—312.

Killworth, P. D., Bernard, R,, & McCarty, C. (1984). Measuring patterns of acquaintanceship. *Current Anthropology*, *25*, 381—392.

Kirke, D. M. (1996). Collecting peer data and delineating peer networks in a complete network. *Social Networks*, *18*, 333—346.

Klovdahl, A. S. (2005). Social network research and human subjects protection: Towards more effective infectious disease control. *Social Networks*, *27*, 119—137.

Knoke, D. (2001). *Changing organizations: Business networks in the new political economy*. Boulder, CO: Westview Press.

Knoke, D. (in press). Playing well together: Creating corporate social capital in strategic alliance networks. *American Behavioral Scientist*.

Knoke, D., & Burt, R. (1983). Prominence. In R. S. Burt & M. J. Miner (Eds.), *Applied network analysis: A methodological introduction* (pp. 195—222). Beverly Hills CA: Sage.

Knoke, D., & Kuklinski, J. H. (1982). *Network analysis*. Beverly Hills, CA: Sage.

Knoke, D., & Laumann, E. O. (1982). The social structure of national policy domains: An exploration of some structural hypotheses. In P. V. Marsden & N. Lin (Eds.), *Social structure and network analysis* (pp. 255—270). Beverly Hills, CA: Sage.

Knoke, D., Pappi, F. U., Broadbent, J., & Tsujinaka, Y. (1996). *Comparing policy networks: Labor politics in the U. S., Germany, and Japan*. New York: Cambridge University Press.

Knoke, D., & Rogers, D. (1979). A blockmodel analysis of interorganizational networks. *Sociology and Social Research*, *64*, 28—52.

Knox, H., Savage, M., & Harvey, P. (2006). Social networks and the study of relations: Networks as method, metaphor and form. *Economy and Society*, *35*, 113—140.

Koehly, L. M., & Pattison, P. (2005). Random graph models for social networks: Multiple relations or multiple raters. In P. J. Carrington, J. Scott, & S. Wasserman(Eds.), *Models and methods in social network analysis*(pp. 162—191). New York: Cambridge University Press.

Kossinets, G. (2003). Effects of missing data in social networks. *Social Networks*, *28*, 247—268.

Krackhardt, D. (1987). Cognitive social structures. *Social Networks*, *9*, 109—134.

Krackhardt, D., & Kilduff, M. (1999). Whether close or far: Perceptions of balance in friendship networks in organizations. *Journal of Personality and Social Psychology*, *76*, 770—782.

Kronenfeld, D. B., & Kronenfeld, J. (1972). Toward a science of design for successful food service. *Institutions and Volume Feeding Management*, *70*, 38—44.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.

Kumbasar, E., Romney, K., & Batchelder, W. (1994). Systematic biases in social perception. *American Journal of Sociology*, *100*, 477—505.

LaPierre, R. T. (1934). Attitudes vs. actions. *Social Forces*, *13*, 230—237.

Laumann, E., Marsden, P. V., & Prensky, D. (1983). The boundary-specification problem in network analysis. In R. Burt & M. Minor(Eds.), *Applied network analysis*(pp. 18—34). Beverly Hills, CA: Sage.

Laumann, E. O., & Knoke, D. (1987). *The organizational state: Social choice in national policy domains*. Madison: University of Wisconsin Press.

Laumann, E. O., Mardsen, P. V., & Prensky, D. (1989). The boundary specification problem in network analysis. In L. C. Freeman, D. R. White, & A. K. Romney(Eds.), *Research methods in social network analysis*(pp. 61—87). Fairfax, VA: George Mason University Press.

Lin, N. (1982). Social resources and instrumental action. In P. V. Marsden & N. Lin(Eds.), *Social structure and network analysis* (pp. 131—145). Beverly Hills, CA: Sage.

Lin, N., & Dumin, M. (1986). Access to occupations through social ties.

Social Networks, *8*, 365—385.

Lin, N., Fu, Y., & Hsung, R.-M. (2001). The position generator: Measurement techniques for investigations of social capital. In N. Lin, K. Cook, & R. Burt (Eds.), *Social capital: Theory and research* (pp. 57—81). Hawthorne, NY: Aldine de Gruyter.

Luce, D., & Perry, A. D. (1949). A method of matrix analysis of group structure. *Psychometrika*, *14*, 95—116.

Marsden, P. (2002). Egocentric and sociocentric measures of network centrality. *Social Networks*, *24*, 407—422.

Marsden, P. V. (1987). Core discussion networks of Americans. *American Sociological Review*, *52*, 122—131.

Marsden, P. V. (1993). The reliability of network density and composition measures. *Social Networks*, *15*(4), 399—421.

Marsden, P. V. (2005). Recent developments in network measurement. In P. Carrington, J. Scott, & S. Wasserman, *Models and methods in social network analysis* (pp. 8—30). New York: Cambridge University Press.

McCarty, C., Bernard, R., Killworth, P., Shelley, G., & Johnsen, E. (1997). Eliciting representative samples of personal networks. *Social Networks*, *19*, 303—323.

McPherson, M., Smith-Lovin, L., & Brashears, M. E. (2006). Social isolation in America: Changes in core discussion networks over two decades. *American Sociological Review*, *71*, 353—375.

Meyer, M. (2000). What is special about patent citations? Differences between scientific and patent citations. *Scientometrics*, *49*(1), 93—123.

Mische, A., & Pattison, P. E. (2000). Composing a civic arena: Publics, projects, and social settings. *Poetics*, *27*, 163—194.

Mitchell, J. C. (1969). *Social networks in urban situations: Analyses of personal relationships in Central African towns*. Manchester, UK: Manchester University Press.

Mohr, J. W. (1998). Measuring meaning structures. *Annual Review of Sociology*, *24*, 345—370.

Mohr, J. W., & Duquenne, V. (1997). The duality of culture and practice: Poverty relief in New York City, 1888—1917. *Theory and Society*, *26*, 305—356.

Mokken, R. J. (1979). Cliques, clubs and clans. *Quantity and Quality*, *13*, 161—173.

Moreno, J. L. (1934). *Who shall survive?* Washington, DC: Nervous and Mental Disease Publishing Company.

Morrissey, J., Tausig, M., & Lindsey, M. (1985). *Network analytic methods for mental health service system research: A comparison of two community support systems (No. 6)*. Washington, DC: National Institute of Mental Health Series BN.

Mouttapa, M., Valente, T. W., & Gallaher, P. (2004). Social network predictors of bullying and victimization. *Adolescence*, *39*, 315—335.

Nadel, S. F. (1957). *The theory of social structure*. London: Cohen and West.

Neuman, W. L. (2000). *Social research methods: Qualitative and quantitative approaches* (4th ed.). Boston: Allyn & Bacon.

Newcomb, T. (1961). *The acquaintance process*. New York: Holt, Rinehart and Winston.

Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, *45*, 167—256.

de Nooy, W., Mrvar, W. A., & Batagelj, V. (2005). *Exploratory social network analysis with Pajek*. New York: Cambridge University Press.

Nowicki, K. & Snijders T. (2001). Estimation and prediction for stochastic blockstructures. *Journal of the American Statistical Association*, *96* (455), 1077—1088.

Nyblom, J., Borgatti, S., Roslakka, J., & Salo, M. A. (2003). Statistical analysis of network data: An application to diffusion of innovation. *Social Networks*, *25*, 175—195.

Padgett, J. F. & Ansell, C. K. (1993). Robust action and the rise of the Medici, 1400—1434. *American Journal of Sociology*, *98*, 1259—1319.

Pampel, F. C. (2000). *Logistic regression: A primer*. Thousand Oaks, CA: Sage.

Pattison, P. (1988). Network models: Some comments on papers in this special issue. *Social Networks*, *10*, 383—411.

Pattison, P., & Wasserman, S. (1999). Logit models and logistic regressions for social networks: Ⅱ. Multivariate relations. *British Journal of Mathematical and Statistical Psychology*, *52*, 169—193.

Peay, E. R. (1980). Connectedness in a general model for valued networks. *Social Networks*, *2*, 385—410.

Perrow, C. (1986). *Complex organizations: A critical essay*. New York: Random House.

Podolny, J. M. & Stuart, T. (1995). A role-based ecology of technological change. *American Journal of Sociology*, *100*, 1224—1260.

Raab, J., & Milward, H. B. (2003). Dark networks as problems. *Journal of Public Administration Research and Theory*, *13*, 413—439.

Rice, R. (1990). Computer-mediated communication system network data: Theoretical concerns and empirical examples. *International Journal of Man-Machine Studies*, *32*, 627—647.

Rice, R. (1994). Network analysis and computer-mediated communication systems. In S. Wasserman & J. Galaskiewicz(Eds.), *Advances in social network analysis*(pp. 167—203). Newbury Park, CA: Sage.

Rice, R., Borgman, C., Bednarski, D., & Hart, P. J. (1989). Journal-to-journal citation data: Issues of reliability and validity. *Scientometrics*, *15*, 257—282.

Robins, G., Pattison, P., & Wasserman, S. (1999). Logit models and logistic regressions for social networks: Ⅲ. Valued relations. *Psychometrika*, *64*, 371—394.

Romney, K., & Weller, S. (1984). Predicting informant accuracy from patterns of recall among individuals. *Social Networks*, *6*, 59—77.

Romney, K., Weller, S., & Batchelder, W. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist*, *88*, 313—338.

Ruan, D. (1998). The content of the General Social Survey discussion networks: An exploration of General Social Survey discussion name generator in a Chinese context. *Social Networks*, *20*, 247—264.

Ruan, D., Freeman, L., Dai, X., Pan, Y., & Zhang, W. (1997). On the changing structure of social networks in urban China. *Social Networks*, *19*, 75—89.

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika*, *31*, 581—603.

Salganik, M. J., & Heckathorn D. D. (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology*, *34*, 193—240.

Schiffman, S., Reynolds, L., & Young, F. (1981). *Introduction to multidimensional scaling*. New York: Academic Press.

Schwartz, J. E. (1977). An examination of CONCOR and related methods for blocking sociometric data. *Sociological Methodology*, *7*, 255—282.

Schweizer, T. (1996). Actor and event orderings across time: Lattice representation and Boolean analysis of political disputes in Chen Village, China. *Social Networks*, *18*, 247—266.

Scott, J. (1991). *Social network analysis: A handbook*. London: Sage.

Scott, J. (2000). *Social network analysis: A handbook* (2nd ed.). Thousand Oaks, CA: Sage.

Seidman, S. B. (1983). Network structure and minimum degree. *Social Networks*, *5*, 269—287.

Simmel, G. (1908). *Sociology: Investigations on the forms of sociation*. Leipzig, Germany: Duncker and Humblot.

Slez, A., & Martin, J. L. (2007). Political action and party formation in the United States constitutional convention. *American Sociological Review*, *72*(1), 42—67.

Smith, D., & Timberlake, M. (2002). Hierarchies of dominance among world cities: A network approach. In S. Sassen (Ed.), *Global networks, linked cities* (pp. 117—141). New York: Routledge.

Snijders, T. A. B., Steglich, C., Schweinberger, M., & Huisman, M. (2007). *Manual for SIENA version 3*. Groningen, Netherlands: University of Groningen, ICS, Department of Sociology.

Stoebenau, K., & Valente, T. W. (2003). Using network analysis to understand community-based programs: A case study from highland Madagascar. *International Family Planning Perspectives*, *29*(4), 167—173.

Stork, D., & Richards, W. (1992). Nonrespondents in communication network studies: Problems and possibilities. *Group & Organization Management*, *17*(2), 193—210.

Strang, G. (1988). *Linear algebra and its applications* (3rd ed.). San Diego: Harcourt, Brace, Jovanovich.

Strauss, D., & Ikeda, M. (1990). Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association*, *85*, 204—212.

Taylor, S., & Fiske, S. (1978). Salience, attention, and attribution: Top-of-the-head phenomena. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, (*vol. 11*) pp. 249—288. New York: Academic Press.

Tenkasi, R. V., & Chesmore, M. C. (2003). Social networks and planned organizational change: The impact of strong network ties on effective change implementation and use. *Journal of Applied Behavioral Science*, *39*, 281—300.

Tsuji, R. (1997). *PermNet, version 0.94*. Sapporo: Hokkaido University.

Tsvetovat, M., & Carley K. M. (2005). Structural knowledge and success of anti-terrorist activity: The downside of structural equivalence. [Electronic version]. *Journal of Social Structure*, *6*(2).

Useem, M. (1979). The social organization of the American business elite and participation of corporation directors in the governance of American institutions. *American Sociological Review*, *44*, 553—572.

Valente, T. (1995). *Network models of the diffusion of innovations*. Cresskill, NJ: Hampton Press.

Van der Gaag, M., & Snijders, T. (2004). Proposals for the measurement of individual social capital. In H. Flap & B. Volker(Eds.), *Creation and returns of social capital: A new research program*(pp. 199—219). New York: Routledge.

Vedres, B., Bruszt, L., & Stark, D. (2004). Organizing technologies: Genre forms of online civic association in Eastern Europe. *Annals of the American Academy of Political and Social Science*, *597*, 171—188.

Wasserman, S. & Faust, K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.

Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p. *Psychometrika*, *61*(3), 401—425.

Wasserman, S., & Robins, G. (2005). An introduction to random graphs, dependence graphs, and p*. In P. J. Carrington, J. Scott, & S. Wasserman(Eds.), *Models and methods in social network analysis* (pp. 148—161). New York: Cambridge University Press.

Watts, D. J. (2003). *Six degrees: The science of a connected age*. New York: Norton.

Wellman, B. (1999). From little boxes to loosely bounded networks: The privatization and domestication of community. In J. L. Abu-Lughod (Ed.), *Sociology for the twenty-first century*(pp. 94—114). Chicago: University of Chicago Press.

Wellman, B., & Berkowitz, S. D. (1988). *Social structures: A network ap-*

proach. Cambridge UK: Cambridge University Press.

White, D. (1996). Statistical entailments and the Galois lattice. *Social Networks*, *18*, 201—215.

White, D., & Reitz, K. (1983). Graph and semigroup homomorphisms on networks of relations. *Social Networks*, *5*, 193—234.

White, H. C., Boorman, S. A., & Breiger, R. L. (1976). Social structure from multiple networks. I: Blockmodels of roles and positions. *American Journal of Sociology*, *81*, 730—780.

White, K., & Watkins, S. C. (2000). Accuracy, stability and reciprocity in informal conversational networks in rural Kenya. *Social Networks*, *22*, 337—355.

White, M., Kim, A., & Glick, J. J. (2005). Mapping social segregation. *Sociological Methods and Research*, *34*, 173—203.

Winship, C., & Mandel, M. (1983). Roles and positions: A critique and extension of the blockmodeling approach. *Sociological Methodology*, *14*, 314—344.

Wu, L. L. (1983). Local blockmodel algebras for analyzing social networks. *Sociological Methodology*, *14*, 272—313.

Yang, S., & Hexmoor, H. (2004). Measuring optimal connections in large networks: A new algorithm and its application. *Journal of Mathematical Sociology*, *28*(3), 197—213.

Yang, S., & Knoke, D. (2001). Optimal connections: Strengths and distance in valued graphs. *Social Networks*, *23*, 285—295.

译名对照表

actors	行为人
alters	行为客体
actor contacts	行为人联系
adjacency matrix	邻接矩阵
advice-seeking network	建议寻求网络
affiliation matrix	隶属矩阵
affiliation networks	隶属网络
Americans' core discussion groups	美国人的核心讨论圈
analysis of lattices	点阵分析
archival documents	档案资料
a single inclusive cluster	单一的全涵盖聚类
asymmetric choice	非对称选择
attributes	特征
authority/power relations	权力关系
automorphic equivalence	自形等价
average link	平均联结
average method	均值法
betweenness	中间性
betweenness centrality	中间性中心度
binary nondirected	二进制无向
binary directed	二进制有向
bipartite graph	二分图
blockmodel	块模型
boundary penetration relations	交叉关系
boundary specification	范围界定
bridge	桥
cascade modules	级联模块
categoric variables	分类变量
centrality	中心度
chain-referral method	连锁推荐法
chaos theory	混沌理论

checklist	提名清单
checklist method	名单核对法
circuit	回路
clique	小团体
clique size	小团体规模
closed walks	封闭行程
closeness	亲密性
closeness centrality	亲密中心度
cluster	群
clustering	聚类
clustering algorithms	聚类演绎法
cognitive maps	认知图
Cognitive Social Structure(CSS)	认知社会结构
cohesion	内聚力
cohesiveness	内聚
cohesive subgroup	内聚型亚群体
communication relations	沟通关系
complementarity hypothesis	异性相吸的互补假设
complete link	完全联结
complete reliability	完全可信
complete networks	全幅网络
complete network measures	全幅网络测量
complex theory	复杂理论
components	组成部分
Computer-Mediated Communication(CMC)	电子化交流
computer-monitored network data	计算机控制网络数据
connection	连接
consensus formula	共识模型
contact	接触
correspondence analysis	对应分析
cutpoints	分界点
degree centrality	程度中心度
dendrogram	树形图

density cutoff	密度临界值
density matrix	密度矩阵
diameters	直径
digraph	有向图
digraph clique	有向图小团体
directed dyadic relations	有向对偶关系
directed graph	有向图表
direct relation	定向关系
dissimilarities	异质性
diversity	多样性
divisiveness	分裂性
descriptive statistics	描述统计学
dispersion	离差
drawn sociograms	手绘社会关系图
dyad	对点
dyadic network	对偶网络
dyadic relation	对偶网络关系
dyadic units	对偶组合
ego	行为主体
ego actor	主体行为人
egocentric analysis	自我中心分析法
egocentric networks	自我中心网络
egocentric network measures	自我中心网络的测量
ego-density	自我密度
eigenvector centrality	特征向量中心度
error term	误差项
Euclidean distance	欧氏距离
event	事件
event-based strategies	基于事件的策略
event contacts	事件联系
expansiveness	扩展程度
faction	派系
first zone	第一关系圈

fixed list approach	固定名单法
fixed list selection	固定名单选择法
frequency	频率
galois connection	伽罗瓦连接
gang	帮派
General Social Survey(GSS)	综合社会调查
geodesic distance	测地距
geographical constraint	地域限制
graphs	图表法
graph theory	图论
group closeness centralization	组亲密中心度
group degree centralization	组程度中心度
hierarchical agglomerative cluster analysis	分层聚类分析
hierarchical clustering	分层聚类法
high degree-seeking strategy	高相关度寻找法
homophily hypothesis	羽禽群居的同质假设
human capital	人力资本
indegrees	节点入度
Index of Qualitative Variation(IQV)	定性变化指数
Individual Differences Scaling(INDSCAL)	个体差异标度
informant bias	信息偏差
initial sample	初始样本
initiators	发起人
instrumental name generator	工具性提名生成法
instrumental relations	利用关系
instrumental ties	工具性连接点
intensity	强度
interlocking roles	互锁角色
internal consistency reliability	内在一致性信度
interobserver reliability	观察者间信度
isolated structures	孤立结构
isomorphism	同质性
isomorphic equivalence	同构等价

isomorphic subgraphs	同构子图
Jaccard's coefficient	Jaccard 系数
jointly dyadic property	共同的对偶特征
K-core method	K 核心法
Kevin Bacon game	凯文·培根游戏
kinship and descent relations	亲属关系
labeled points	标示点
lattices	点阵
longitudinal information	纵向信息
matrices	矩阵法
maximal complete subgraph	最大完全子图
mean egocentric density measure	自我中心网络密度测量均值
mean sex diversity	性别多样化均值
measurement of relations	关系测量
missing data	数据缺失
MultiDimensional Scaling(MDS)	多维标度
multiple-name generator methods	多个提名生成法
multiple networks	多重网络
multiple social support name generator	多重社会资源提名法
mutualchoice	双向选择
name generators	提名法
name interpreters	姓名阐释法
National Longitudinal Study of Adolescent Health	美国青少年健康纵向调查
network position measures	网络位置测量
network sampling	网络抽样
network size	网络规模
network structure	网络结构
nodes	节点
nodal degree	节点度
nodal indegree	节点入度
nodal outdegree	节点出度
nodal response rate	节点应答率

nominalist strategies	唯名论界定策略
nondirected graph	非定向图
nondirected dyadic relations	非定向的对偶关系
nonoverlapping clusters	非交叉聚类
nonreciprocal relations	单向关系
normalized actor degree centrality	标准化行为人程度中心度
numerical constraint	数量限制
one-mode network	一模网络
online social networking	在线网络交流
ordinary least square regression methods	最小二乘回归法
pair characteristics	二元特性
paths	路径
Pearson's correlation coefficient	皮尔逊相关系数
points	点
popularity	受欢迎程度
positon	位置
positional generators	位置选择法
positional strategies	定位策略
prestige	声望
realist strategies	唯实论界定策略
receivers	接收者
receptivity	接纳程度
regular equivalence	正规等价
relational response rate	关系应答率
relational strategies	关系策略
relationship measures	关系的测量方法
reliability	信度
reputational method	声望法
resource generators	资源选择法
respondent-driven sampling	应答驱动抽样
Reverse Small-World method(RSW)	反向小世界方法
role/content constraint	角色/内容限制
sampling bias	抽样偏差

sampling frame	抽样框
sampling units	抽样单位
saturated structures	饱和结构
seciometry	社会计量法
self-relation	自我关系
self-report questionaire surveys	自填式问卷调查
sender-receiver	发送者—接收者
sentiment relations	情感关系
similarity	相似性
single link	单个联结
single-name generator methods	单个提名生成法
singleton cluster	单独的聚类
Singular Value Decomposition(SVD)	奇异值分解
six degree of separation	六度分隔
small-world idea	小世界理论
small-world phenomena	小世界现象
snowball sampling	滚雪球抽样法
social capital	社会资本
social cognitive structure	认知社会结构
social identity approaches	社会认同途径
social resources theory	社会资源理论
social support networks	社会支持网络
sociogram	社会关系图
sociomatrix	社会矩阵
sociometric	社会计量学
sociometric techniques	社会计量方法
source-target	源头—目标
split-half reliability	折半信度
standard deviation	标准差
standard deviation for continuous variables	连续变量的标准差
star graph	星形图
strength	力度
stress	应力

strongly connected	强连接
structural equivalence	结构等价
structural relations	结构关系
structurally equivalent actors	结构等价的行为人
subgraph	子图
subgroup	亚群体
substantialist approaches	实体论方法
subsystem	子系统
summation method	求和法
temporal constraint	时间限制
terminators	终结者
test-retest reliability	再测法信度
total network	全部网络
transaction relations	交换关系
tree diagram	树图
triadic relations	三元关系
two-mode network	二模网络
type of tie	连接模式
unilaterally connected	单向连接
utility	效用
valued directed graph	赋值有向图
valued graph	赋值图
valued nondirected graph	赋值无向图
visual displays	视觉演示
weak connected	弱连接
Weighted MultiDimensional Scaling(WMDS)	加权多维标度



Social Network Analysis, 2nd Edition

上海市版权局著作权合同登记号:图字 09-2009-546